Assembly responses of hippocampal CA1 place cells predict learned behavior in goal-directed spatial tasks on the radial eight-arm maze

Haibing Xu[‡], Peter Baracskay[#], Joseph O'Neill[*†], Jozsef Csicsvari[*]

Institute of Science and Technology Austria (IST Austria)

Am Campus 1, Klosterneuburg, A – 3400, Austria

Lead Contact: Jozsef Csicsvari

[*]Correspondence: jozsef.csicsvari@ist.ac.at, ONeillJ9@cardiff.ac.uk

[‡]Current address: Kirby Neurobiology Center, Harvard Medical School, Boston, MA 02115, USA.

[#]Current address: Department of Biochemistry, Institute of Biology, Eötvös Loránd University, Budapest, H-1117 Hungary

[†]Current address: School of Psychology, Cardiff University, Cardiff CF10 3AT, UK

Summary

Hippocampus is needed for both, spatial working and reference memories. Here, using a radial eight-arm maze, we examined how the combined demand on these memories influenced CA1 place cell assemblies while reference memories were partially updated. This was contrasted with control tasks requiring only working memory or the update of reference memory. Reference memory update led to the reward-directed place field shifts at newly-rewarded arms and to the gradual strengthening of firing in passes between newly-rewarded arms but not between those passes that included a familiar-rewarded arm. At the maze center, transient network synchronization periods preferentially replayed trajectories of the next chosen arm in reference memory tasks, but the previously visited arm in the working memory task. Hence, reference memory demand was uniquely associated with a gradual, goal novelty-related reorganization of place cell assemblies and with trajectory replay that reflected the animal's decision of which arm to visit next.

Introduction

The hippocampus plays a critical role in navigation and spatial learning, underpinned by place coding by its principal cells (Morris et al., 1982; O'Keefe and Dostrovsky, 1971). Collectively, ensembles of place cells in the hippocampus encode a cognitive map that is uniquely configured for each environment explored (Muller and Kubie, 1987; O'Keefe and Nadel, 1978). As a result, the instantaneous activity of place cell assemblies provides upstream brain regions with information regarding where the animal is within a specific context, at any given moment in time. However, place cell firing encodes more than locations: neuronal firing patterns can be dynamically modulated when the animal is engaged in spatial learning tasks (Frank et al., 2000; Hok et al., 2007; Markus et al., 1995; Wood et al., 2000). Such changes in firing patterns are correlated with task performance and likely represent mechanisms that support memory (Allen et al., 2012; Dupret et al., 2010; Ferbinteanu and Shapiro, 2003).

The hippocampus is required to form long-term memories of salient places, as well as to flexibly use task rules in order to locate a reward from a discrete number of possible locations (Morris et al., 1982; Olton and Samuelson, 1976; Rawlins and Olton, 1982). These distinct cognitive tasks are accompanied by different task-related changes in place cell activity. Goal locations are disproportionally represented by place cells (Hok et al., 2007; Hollup et al., 2001). Moreover, when new goal locations are learned in a previously familiar environment, many hippocampal place cells remap their place fields near to the novel goal locations (Dupret et al., 2010). The degree of overrepresentation at a goal reflects memory performance, indicating a role in learning.

The second category of hippocampal-dependent spatial tasks requires the animal to infer a future reward location by its previous spatial experience and the rules of the task itself. Under these conditions, place cells tend to maintain similar spatial selectivity across trials and, instead, exhibit firing rate changes that correlate with task parameters (Ainge et al., 2012; Allen et al., 2012; O'Neill et al., 2017). While running on a maze-arm before a choice point, place cell firing is influenced by the future and past arm choices of the animal, as well as by trial stages such as the encoding or recall stages of the task (Ferbinteanu and Shapiro, 2003; Frank et al., 2000; Spellman et al., 2015; Wood et al., 2000). This task-contingent rate remapping may be related to the decision making during the working memory task. Alternatively, this may merely encode the episodic-like experience of the animal performing the task, information that could then be used to select the appropriate goal location (Eichenbaum et al., 1999). Indeed, such episodic-like coding is suggested by the rate remapping of these place cells between different trial stages (Leutgeb et al., 2004). In summary, goal-related remapping has been examined in relation to goal learning in open environments whereas task-contingent rate remapping has been typically examined in maze environments, in tasks requiring working memory. It is unclear however how

these coding schemes may interact with each other when new goals are learned and at the same time animals have to use working memory.

In addition to the rate remapping of individual place cells, spatial memory is also likely supported by place cell assembly firing patterns during brief periods of synchronous firing. Activity during such high synchrony events (HSE) often encodes a temporally-compressed replay of movement trajectories of an explored environment (Davidson et al., 2009; Foster and Wilson, 2006; Lee and Wilson, 2002). Several studies have examined such trajectory replay during memory tasks by detecting replay events in both theta epochs and hippocampal sharp wave ripples (SWR) or SWR-associated HSEs. The involvement of SWR-associated replay in spatial working memory is suggested by the observation that blocking these SWRs cause working memory deficits (Jadhav et al., 2012). Furthermore, the SWR response of cells can predict whether the animal will make an error; however, the reactivated trajectories did not predict the behavioral choice of the animal (Singer et al., 2013).

Similarly, theta oscillation-associated replay did not predict which of the alternative goals are selected next; both could be replayed (Johnson and Redish, 2007). Since neither theta nor SWRs-associated replay predicted the future choice, it is suggested that the replayed choice alternatives might help the animal to decide about its next choice (Redish, 2016). However, in these studies only two spatial choices were offered, therefore, it is not clear whether replay reflects only the viable alternatives or reflects merely the experience of the animal running on the maze, without replaying the viable alternatives only.

However, in some instances replay may predict parameters related to the spatial choice of the animal. While animals run on the same circular path, the length of the replayed trajectories during theta oscillations can indicate the animal's choice between near or distant goals (Wikenheiser and Redish, 2015). Moreover, the direction of the replayed trajectories during HSEs correlated with the direction of the goal locations relative to the animal's position (Pfeiffer and Foster, 2013). One possible explanation for these differences is that waking trajectory replay might play a direct role in spatial decisions when an allocentric strategy is used to navigate to fixed goals. However, in other tasks that require spatial task-contingent decisions, replay may represent viable alternatives, assisting other brain regions in decision making.

To test how different neuronal coding schemes that reflect spatial memory traces interact with each other during a combined working and a longer-term memory demand, and how SWR-associated replay might be involved in decision making, we recorded while rats solved a radial eight-arm maze task. This task required navigation to fixed goal locations and keeping track of the visited arms at the same time, thus, requiring the use of both spatial reference and working memory. This was compared to that of two other task variants, requiring only working or reference memory. We found that reference memory demand influenced place cell assembly coding in several ways, which were not seen during the task requiring pure working memory. Our results also suggest that sequence replay plays a role in decision making in tasks requiring reference memory. In contrast, replay during the pure working memory task was not associated

with the future choice or viable choice alternatives; instead, the past choice was preferentially replayed.

Results

We recorded CA1 multiple-unit activity in four animals while they performed a task on a radial eight-arm maze, in which they had to retrieve food rewards at the end of three arms (Figure 1A). On each recording day, one rewarded arm remained constant from the previous day (familiar-goal arm) whereas the remaining two goal arms were selected from those that did not contain rewards the day before (novel-goal arms). Therefore, in each recording day animals had to partially update their goal-related reference memory representations to remember which arms contained food. In each trial, they also had to keep track of the previously visited arms. In addition to this combined working and reference memory task ('combined task'), we recorded during two control tasks in three further animals, using the same familiar and novel goal arm selection, where the animal was required to use either spatial working or reference memory only. In the spatial reference memory task, the animal was prevented from making working memory errors by blocking the visited arms in a trial, whereas in the spatial working memory task, arms without food were blocked. Finally, in three additional animals, we repeated the working and reference memory experiments in which the order of training of the two tasks was reversed.

Animals took 10-15 trials to learn the combined task, while in the control tasks fewer trials (5-10 trials) were required (Figure 1B). During learning, animals made more reference and working memory errors in the combined task than in the other tasks, but once the new goals were learned, the performance was similar (Figure 1C-D).

CA1 place cells move their place fields towards novel goal locations even in a familiar environment (Dupret et al., 2010). In order to assess changes in place coding, we first subdivided learning into five blocks of six trials and established place maps in each block. In the combined task, the place field similarity between learning trial blocks and the subsequent post-probe session increased as learning progressed, indicating that place maps reorganized gradually during learning ($r=0.6885$, $P<0.0001$, Figure 2A-B). A similar reorganization was seen in the reference memory task **(**$r=0.6994$, $P<0.0001$, Figure S1A-B**)** but not in the working memory task **(**$r=0.1671$, $P=0.3028$, Figure S1D,E**)**. The familiar-goal arm place fields showed minimal reorganization during all tasks (all $r<0.19$, $p>0.14$, Figure 2C, Figure S1C,F), while the novel-goal arms exhibited gradual changes for the combined and reference memory tasks **(**all $r>0.52$, $p<0.0001$**)**. Place cells representing the familiar-goal arm may not have shifted their place fields because they had already undergone a place field shift during the previous learning day. Differences in the organization of place fields between the novel and familiar arms did not occur as a result of different running speed. While the speed of the animal increased slightly in later trials, overall speed was similar both on the familiar and novel arms (Figure S2 A, top).

Moreover, the typical running speed at each point on the arm was similar when novel and familiar arms were compared (Figure S2 A, bottom).

Many place cells fired unidirectionally on the arm and the proportion of these cells increased across training days (Figure 2D), in agreement with past work showing such increase with path familiarity (Battaglia et al., 2004; Markus et al., 1995; Navratilova et al., 2012). By plotting the place fields across trials, we observed their gradual shift towards the goal locations, but only during reward-bound passes (i.e., runs from the center towards the goal at the end of the arm) of novel-goal arms and not in center-bound passes (i.e., during runs from the reward location at the end of the arm to the center of the maze, Figure 3A). Place fields in novel reward-bound passes exhibited a significant shift ($P<0.001$, binomial test, Figure 3B), increasing the proportion of cells firing near the rewards at the end of learning compared to the beginning ($P<0.001$, one-tailed K-S test, Figure 3C). Moreover, a similar shift was seen during the spatial reference but not in the spatial working memory task **(**Figure S3**)**. Some of these place cells only started to fire after the first trial. Indeed, the number of cells not firing in the first trial was higher on the novel- than familiar-goal arms and fewer cells showed such behavior in the working memory task (Figure S2 B-C). Past work has also seen the emergence of new place fields during running of familiar tracks (Mehta and Mc Naughton, 1997; Monaco et al., 2014). Thus, the formation of new reference memories was accompanied by a gradual shift of place field firing specifically in goal-bound trajectories, while spatial firing patterns remained stable in those goal-arms learned the previous day. Since the working memory task showed no such changes in place selectivity, we next tested whether firing rates were modulated during the task and whether such rate changes reflected working or reference memory task demands.

In maze tasks, the firing rates of place cells are often modulated by past or future arm choices of the animal (Ferbinteanu and Shapiro, 2003; Frank et al., 2000; Wood et al., 2000). Here, we examined whether the arm the animal came from influenced its reward-bound firing and whether its center-bound firing predicted its next arm choice (Figure 4). We found that each session contained an average of 42.6±2.2% (mean ± SEM, n=16) of reward-bound firing place cells and 41.8±2.1% (n=16) of the center-bound firing cells that exhibited a significant rate modulation independent of speed ($p<0.05$, ANCOVA see Start Methods). A similar proportion of cells showed rate modulation in each of the other tasks (Reference memory: Reward-bound 43.3±3.4%, Center-bound 40±4.1%; Working memory: Reward-bound 41.5±6%, Center-bound 43.2±3.9%; mean ± SEM, n=8). Surprisingly, in the combined and reference memory tasks, the firing rate modulation reorganized towards the end of learning and exhibited a modulation by prospective and retrospective goal novelty (Figure 4A-B). Retrospective goal novelty-related firing was demonstrated by the observation that in the second half of learning, but not in the first half, the reward-bound firing rate of cells in a novel-goal arm was higher if the animal entered it from the other novel-goal arm as compared to the familiar-goal arm.

Similarly, prospective goal novelty-related firing was seen in center-bound passes on a novel goal-arm. In this case, firing was stronger when the animal went next to the other novel-goal

arm as compared to the familiar-goal arm. Both of these effects were seen in the second half of the learning trials but not in the first half. To quantify these effects, the firing rate of cells in an arm was compared in passes in which the other arm of the trajectory was a familiar- or a novel-goal arm and the normalized rate difference (i.e., difference of rates divided by their sums) was calculated (Figure 4C, Figure S4 top histograms). During the combined and reference memory tasks the rate score distribution of cells showed a positive bias (all $P<0.001$, binomial test) for the second half of learning but not for the first half (all $P>0.64$, binomial test) indicating that rates were higher in novel-novel than in novel-familiar passes. This effect was absent in the working memory task (all $P>0.56$, binomial test). Similar results were obtained when considering only those cells that exhibited significant rate modulation between that familiar and novel goal arms independent of speed (cells with $p<0.05$ in the ANCOVA described above, Figure 4D, Figure S4 bottom histograms). Hence prospective and retrospective goal novelty-related firing was seen independent of the speed of the animal.

We have observed conditional firing depending on the past and future paths in all of our tasks. However, the prospective and retrospective goal novelty-dependent firing that developed during learning was observed only in the combined and reference memory tasks. Next, we examined whether place cell assembly activity always encoded the current location of the animal or, on occasions, reflected entire replayed trajectories on the maze. Such trajectory replay may be related to working memory demand because the blockade of waking SWRs resulted in spatial working but not reference memory deficits on the W-maze task (Jadhav et al., 2012). However, these may still reflect goal-related trajectories (Pfeiffer and Foster, 2013).

We detected HSEs in which many cells fired action potentials together (See Start Methods, Davidson et al., 2009). The order by which cells fired during HSEs reflected the firing order of place cells during traversal of a goal-arm (Figure 5A). However, the cell order during HSEs might be either similar (i.e., forward order) or reversed relative to the running sequences (Davidson et al., 2009; Diba and Buzsaki, 2007; Foster and Wilson, 2006) and may reflect the most recently explored locations, or future trajectories. We used a Bayesian prediction method to quantify whether these HSE sequences reflected the movement trajectories of the animal (Davidson et al., 2009; Zhang et al., 1998). We divided HSEs into 20ms time windows and tallied the Bayesian probabilities at different locations of the maze, as calculated according to spike counts of different place cells in a given time window (Figure 5B-C). Separate probability maps were calculated for reward-bound and center-bound runs of each goal arm to identify the arm represented by the assembly. To do so, we calculated the trajectory represented within the HSE with the maximum summed probabilities, allowing for both forward and reverse trajectories on reward-bound or center-bound passes (Figure 5D, see Start Methods). These reconstructed trajectories were not random (Figure 6A). We rotated the place field of each active cell within an HSE to see whether replay scores of an event (i.e., mean trajectory probabilities) can occur with the same spike patterns when using decorrelated place maps (Grosmark and Buzsáki, 2016). Our trajectory detection procedure generated significantly weaker probabilities when reconstructed with the rotated place field probability maps than for the original maps (all

P<0.0000001, one-tailed K-S test). In addition, the z-score of the reactivated trajectory probabilities, normalized relative to the mean and SD of the corresponding shuffled distributions, were significantly larger than zero [all P<0.001, binomial test z-scores (mean ± SEM): combined task 1.17±0.068, n=1200; reference task 1.15±0.085, n=605; working task 1.34±0.08, n=691]. These detected HSEs representing reactivated trajectories, which tended to occur during SWRs (Figure 6B), as seen in previous studies (Davidson et al., 2009; Foster and Wilson, 2006; Lee and Wilson, 2002; O'Neill et al., 2017; Pfeiffer and Foster, 2013).

Next, we tested whether these replayed trajectories predicted the future behavior of the animal (Pfeiffer and Foster, 2013; Wikenheiser and Redish, 2015). We separately analyzed HSEs that occurred at goal locations, or in the center of the maze (Figure 7A-B). Trajectory encoding during HSEs that occurred at goal-locations preferentially encoded the current arm in all tasks (all P <0.0001, ANOVA, Figure 7B). However, trajectory encoding in the central area was different between the tasks. In the combined and the reference memory tasks reconstructed trajectories preferentially encoded trajectories of the next visited arm (all P <0.0001, ANOVA, Figure 7A) whereas, in the working memory task, central area HSEs tended to encode the previously visited arm (all P <0.0001, ANOVA).

Replay on the maze can reflect reward-bound or center-bound trajectories and trajectories can reflect places of the maze at the same (i.e., forward) or reverse order relative to how they occurred during behavior (Figure 7B). Therefore, we examined the frequency how often forward or reverse relay encoding goal- or center-bound trajectories occurred. In all tasks, HSEs that occurred at the goal locations tended to exhibit replay of the current arm in the center-bound direction, but in reverse temporal order (all P<0.0000001, Three-way ANOVA with Tukey Post hoc, Figure 7C). However, we again saw significant differences in replay between the tasks in the central area. In the combined and reference memory tasks, replay preferentially encoded reward-bound passes of the next visited arm but in forward order. In contrast, central area HSEs in the working memory task tended to reflect reverse replay, which encoded center-bound trajectories of the previous arm (all P<0.0000001, Three-way ANOVA with Tukey Post hoc, Figure 7C).

In these original experiments involving working and reference memory tasks, training and recordings of the working memory task were performed first followed by the reference memory task. Because the training experience may have influenced the results, we performed experiments in additional three animals in which reference memory training and recordings were performed first. However, the order of the training of the two tasks did not influence our results and we gained identical results (Figure S5). Thus, independent of training order, in the tasks with reference memory demands, replay during HSEs in the center of the maze reflected a forward replay of runs directed towards the future goal, while, in the working memory task, such replay reflected the previous center-bound run, in reverse order.

To test the possible influence of replay events in decision making, we compared replay in passes when the correct arm was chosen (correct pass) and when the animal made an error

(error pass). In the combined and reference memory tasks we used probability maps that encoded all eight-arms to be able to encode the replay of arm-choices associated with reference memory errors. The quality of encoding of trajectory was similar for HSEs occurring during correct or error passes when the z-scored Bayesian trajectory probabilities (relative to the shuffled results) were compared (all $P > 0.5433$, two-way ANOVA). However, in error passes, neither the future nor the past chosen arm trajectories (in the form of a center- or reward-bound trajectory representing forwards or reverse replay) were preferentially encoded (all $P>0.112067$, three-way ANOVA with Tukey Post hoc, Figure 8A). Therefore, the preferential replay of the past or future arms was disrupted in error trials. Next, to infer the rat's decision stage during which HSE events might have occurred, we quantified where replay occurred as the animal run through the central area of the maze. We found that the majority of events occurred right after the animal entered the maze in the correct trials (Figure 8B). This tendency was not seen in error passes where replay HSE events occurred along the entire central area path.

Replay content during active waking periods is influenced by the location and the orientation of the animal (Davidson et al., 2009; Karlsson and Frank, 2009). However, in our case, only about 25-30% of the HSE occurred when the animal faced towards (±30°) the next arm (Figure S6). In the combined and reference memory tasks HSEs preferentially encoded the future arm choice even when the animal faced towards the neighboring arms, up to ±120° away from the next arm while, in working memory task, the previous arm was preferentially encoded in these cases (all $P<0.012$, two-way ANOVA with Tukey Post hoc, Figure S6). The future or past arm was not preferentially encoded for angles >120° (all $P > 0.7078$) or in error trials (all $P > 0.4112$).

The number of HSEs in the central area of the maze might provide additional indications about decision making. We quantified how many HSEs occurred in each trial and at each choice as the animal was crossing the central area (Figure S7). As compared to the seconds half of learning, more replay events occurred during the first half (all $P < 0.016$, ANOVA, Figure S7A-B). Single replay events were more frequent than multiple replay events. During single replay HSE events, the future (combined and reference memory) or the past (working memory) arm was preferentially encoded. In contrast, when multiple HSEs occurred at a choice, no single arm was preferentially encoded, albeit the same arm was encoded more frequently than different arms in such cases (Figure S7 C-D). Moreover, the propensity of multiple events occurring during a choice was higher in error passes than in the correct passes (Figure S7E, all $P < 0.02$, two-way ANOVA with Tukey Post hoc).

Finally, we examined to what degree working memory load influenced replay. We compared replay at different stages of the task starting when the animal made the first choice to the stage when it returned from the last arm (Figure S8). In the correct trials, results were similar as shown for all trials with the exception of two cases. Firstly, in the absence of working memory demand in the first choice, even in the working memory task the future arm was encoded. Secondly, when the animal returned to the center at the end of a trial, without further working

or reference memory demands the past choice of the animal was encoded in all tasks. In error passes no such preference was seen. Therefore, in the absence of working memory demand the future choice was encoded in even in the working memory task while in the cases when the animal did not make further choices the past arm was encoded in all three tasks.

Discussion

In our study, we examined how place cell activity reorganized during a task in which animals had to update their spatial reference memory representations and, at the same time, use spatial working memory. This was compared to neuronal coding during two task variants, requiring either spatial reference or working memory. We have identified several distinguishing firing features of place cells that were linked to the reference memory update. Firstly, place cells gradually moved their place fields towards novel reward locations during learning. Secondly, place cells developed modulated firing rates at an arm depending on prospective or retrospective goal novelty. In the spatial working memory task, we did not find these effects. Furthermore, in tasks requiring spatial reference memories, we found that trajectory replay in the maze center predicted the future arm choice of the animal, with replay proceeding in a forward direction. In contrast, in the spatial working memory task HSEs preferentially encoded the past choice and in a reverse order. However, in all three tasks, replay events at reward locations preferentially encoded the animals' next path to the center, but in reverse temporal order. Thus, during the working memory task, replay typically preceded in a reverse order, which formed representations of the currently traversed or immediately exited arm. When reference memory was required, replay switched modes between the reward location and choice point, from reverse order current arm to the forward replay of the future arm, respectively. Our data point to a role for replay in decision-making because the preferred replay mode in the central area of the maze was disrupted when the animal made an error, in all three task-variants. However, the differences in the type of replay between working and reference memory imply that they can support these memory processes through different systems-level mechanisms.

Goal update-related remapping of place cells

Several studies have demonstrated that many place cells fire near learned goal locations (Hok et al., 2007; Hollup et al., 2001). When these goal locations change, they remap their place fields to represent the new goal configuration, but only when local cues do not mark the goals (Dupret et al., 2010). In agreement with this, we observed the remapping of some of the place fields but only in tasks requiring reference memory update. In our study learning was slow during reference memory update (10-15 trials) and remapping was gradual: several place cells started to fire on the novel-goal arms only after a few trials and they gradually moved their place fields towards the goals. However, place fields in the familiar-goal arm did not remap. Such fields presumably moved the previous day, when the arm was novel. Therefore, the partial update of reference memory representations did not lead to the full reorganization of the hippocampal cognitive map, only the part of the map relating to the novel-goals was updated.

Changes in dopamine levels within the hippocampus may provide a possible mechanism for the observed place field shift. Given that the firing of dopamine cells predicts novel rewards and their firing is increased as animals approach a reward location (Howe et al., 2013; Mirenowicz and Schultz, 1996), ventral tegmental area (VTA) dopamine cells may be active on the novel-goal arms with increasing intensity towards the goal. Indeed, stimulating dopamine fibers originating from the VTA during learning enhance subsequent recall of complex maze paths suggesting a role for dopamine in spatial learning (McNamara et al., 2014). Moreover, VTA dopamine cell stimulation can trigger a shift of place fields towards the location of the stimulation (Mamad et al., 2017).

It is unlikely that this effect was due to place field expansion (Mehta et al., 1997). During place field expansion, place fields shift backward relative to the movement direction of the animal across subsequent trials on a track while in our case forward shift was seen. However, forward place field shift has been seen during the first trials of a spatial alternation task towards reward locations (Lee et al., 2006). This has been observed towards familiar goals without an update of spatial reference memories. However, we only saw such a shift on the novel-goal arms and not in the working memory task, which would be more equivalent to the T-maze alternation task.

Conditional firing is modulated by prospective or retrospective goal novelty

Hippocampal place cells modulate their firing depending on trial types in a variety of spatial maze tasks (Allen et al., 2012; Ferbinteanu and Shapiro, 2003; Frank et al., 2000; Wood et al., 2000). We have observed a similar modulation in all three task variants: the reward-bound firing of place cells was related to the previous arm visit and center-bound firing correlated with the next arm choice. However, this trial type-dependent rate modulation reorganized when reference memories were updated: in later trials, the firing rate of the cells was modulated by goal novelty in the previous or next arms. Moreover, this goal novelty-related firing rate relationship was independent of specific arms considering that this relationship was not present in the beginning of the trial.

Hippocampal pyramidal cells exhibit higher firing rates during the exploration of novel than familiar environments (Csicsvari et al., 2007; Nitz and McNaughton, 1999). The increased pyramidal firing in novel environments may be caused by the increased levels of nonspecific neurotransmitters such as acetylcholine or dopamine in the hippocampus. Indeed VTA cells increase their firing during environmental novelty (McNamara et al., 2014). Moreover, acetylcholine levels are also higher in novel environments, which can also cause pyramidal cell rate changes due to muscarinic receptor modulation (Hasselmo, 1999; Pepeu and Giovannini, 2004). A recent study showed an increased activation of locus coeruleus (LC) neurons in novel environments (Takeuchi et al., 2016). Paradoxically, the blockade of dopamine and not noradrenaline disrupted a novelty-related learning enhancement suggesting the LC axons may release dopamine as well in the hippocampus.

In our data the conditional, goal novelty-related firing increase was seen during both reward-bound and center-bound runs, whereas the place field shift only occurred during reward-bound passes. This suggests that the goal-novelty-related rate modulation and place field shift are linked to different mechanisms, perhaps different subcortical neurotransmitters. For example, dopamine from VTA may cause the place field shift while acetylcholine and/or LC activation may cause the increased firing related to novel goal locations.

Novel goal arm-related rate modulation or place field shift was not seen in the working memory task. Although animals did not have to remember the location of rewarded arms, the working memory errors certainly indicate that animals were influenced by the new task configuration involving partially updated arms. Yet, goal shift or novel-goal arms-related rate changes were not seen here suggesting that these occur only during the update of reference memories.

Trajectory reactivation predicting behavioral choice

Several studies have examined whether trajectory replay predicts the future behavioral choice of the animal. Early studies observed the reactivation of both alternative trajectories at the choice point of a T-maze during theta oscillations (Johnson and Redish, 2007). This has been observed in animals which were hesitant at the choice points and stopped to decide which arm to take next. Theta sequence replay was examined further in a task in which animals had to decide between reward locations with varying reward amount and delays on a circular track (Wikenheiser and Redish, 2015). In such cases, theta sequences replayed longer trajectories when a more distant goal was selected.

Further work suggested that perhaps trajectory replay during SWR-associated HSEs and not theta are important for decision making. This was first shown by revealing differences in SWR firing in error and correct trials on a spatial task on a W-maze (Singer et al., 2013). However, the replayed trajectories did not predict the behavioral choice of the animal. The blockade of SWRs on the same W-maze task caused working but not reference memory deficits suggesting that SWR-associated trajectory replay may be involved in working memory recall (Jadhav et al., 2012). Moreover, in another goal selection task, more SWRs occurred at the previous and in the incoming reward locations when animals did not stop at a decision point (Papale et al., 2016). Therefore, in some tasks, SWRs at reward locations may be even involved in the decision to reach the next goal.

Even in two-dimensional environments, SWR-associated replay may play a role in goal selection: it was shown that heading of the reactivated trajectories predicted the direction of the optimal path to reach a fixed goal (Pfeiffer and Foster, 2013). In our case, the radial eight-arm maze offered unique place cell sequences for each goal, enabling us to identify which arm and thus which goal is to be reached by a reactivated trajectory. At the central maze area, sequence replay predicted the future arm choice in tasks requiring spatial reference memories. These reactivated trajectories preferentially encoded the next reward-bound trajectory of the animal in the forward temporal order. Reactivated trajectories at the reward locations also

encoded the immediate next movement of the animal, this time its return to the center, but in a reverse temporal order. Moreover, in error passes, replay content no longer preferably predicted the behavioral choice of the animal indicating that hippocampal CA1 replay is important for the correct behavioral choice of the animal. In tasks needing reference memory the failure to predict to future choice that the animal used a hippocampus-independent strategy to choose error passes. Also, the increased incidence of multiple replay HSEs and the fact the HSEs did not occur right when the animal entered central area of the maze in error trials can both indicate hesitation similar to other behavioral paradigms where multiple replay events occurred (Johnson and Redish, 2007). In addition, long experience of the animal in navigating on the maze was not required for choice-dependent replay: future arm was replayed both in the combined task as well as in the reference memory task in which the animal was trained to perform this tasks first. Taken together, our finding suggests that sequence replay at the decision point guides the animals' future choice when spatial reference memory is needed for the task.

Why did some of the past work not see the prediction of the future arm choices during replay? In some of these work, two-choice tasks were used involving simple mazes in which the reference memory demand was low. For example, the animals may have learned sequences of left and right turns to complete a trial in such cases. However, it is possible that even in allocentric navigation, depending on experience and training methods, animals may also have to rehearse alternative trajectories before a decision is made.

In the central maze area of the working memory task, the previous trajectory was preferentially replayed in a reverse temporal order, mirroring the replay at the reward locations at the end of the arm. A general rule emerges here applicable for all three tasks: trajectories encoding future goals were encoded in a forward temporal order, while those encoding current and previous goals were represented in a reverse order. Also, in all cases, the part of the trajectory which was closer to the goal was reactivated later in an HSE, while places further away from the goal were reactivated earlier. Manipulating the amount of reward can change the propensity of reverse replay but not forward replay (Ambrose et al., 2016). Perhaps our reward amounts were 'large' so that reverse replay dominated near the goal locations. Nevertheless, we are not aware of previous work highlighting that reward locations tend to occur at the end of a reactivated trajectory, perhaps because in most tasks the end of a path always contains rewards.

In our working memory task, we found that replay during HSEs preferentially replayed the past choice of the animal, not the future choice or choice alternatives, which alone would provide insufficient information for the animal to decide which arm to select next. How can our findings be explained in the light of the SWR blockade experiments that caused working but not reference memory errors on the W-maze (Jadhav et al., 2012)? The W-maze running task required both working and reference memory and in this case our replay indeed predicted the future behavioral choice of the animal. It is likely that in our working memory task the

hippocampus was not involved in arm selection, it only signaled the visited arms to another region (likely the prefrontal cortex), which may be involved in keeping track of the visited arms. Indeed, during a delayed non-match to sample task on the T-maze, the blocking of hippocampal fibers to the medial prefrontal cortex caused working memory deficits during the encoding but not the recall stage (Spellman et al., 2015). Surprisingly, in the combined task requiring both working and reference memory load, replay responses closely mirrored that of the reference memory task. Therefore, the hippocampal replay can participate in two different antagonistic modes while animals solve a spatial memory task: one mode related to reference memory requirement in which the future arm choice is directly encoded. The other mode is associated with pure working memory requirement in which replay may mirror recent behavioral choices or the current arm to communicate them to other brain regions.

This study examined replay responses in the CA1 region of the hippocampus, which provides the primary output of the hippocampus. However, network responses in other subfields such as the CA3 region may have complementary roles in shaping the CA1 output responses of cells. In a recent study, during locally generated CA3 high-frequency oscillations ('CA3 SWRs'), place cells encoded the yet unvisited arms of the maze fired stronger than those representing the visited arms in a complex spatial working memory task (Sasaki et al., 2018). However, it is unlikely that those localized CA3 SWRs encoded specific maze trajectories considering the lack of specificity of place cells to encode specific arms such as the immediate next arm. Additionally, Schaffer collateral synapses are largely suppressed during waking activity (Hasselmo, 1999), leading to the suppression of CA3 excitation of CA1 pyramidal cells during waking SWRs, as shown by reduction of the sharp wave amplitude (O'Neill et al., 2006). In such a case, CA3 cells may influence CA1 pyramidal cells indirectly through recruiting CA1 interneurons, leaving the medial entorhinal cortex to more directly influence CA1 replay (Yamamoto and Tonegawa, 2017). It is possible that the activation of CA3 cells encoding the yet unvisited arms may suppress the firing of similar cells in CA1 during the working memory task used in the Sasaki et al., 2018 study.

In summary, our findings suggest that the hippocampus may be involved in goal selection through sequence replay when alternative goal locations are needed to be localized during the decision, i.e., when reference memory is needed. In contrast, in agreement with past work (Spellman et al., 2015), during pure spatial working memory tasks the hippocampus may only signal the recently visited arm to the prefrontal cortex, which in turn keeps track of arm visits within a given trial, in order to guide future goal selection. In this way, during working memory HSEs and associated SWRs may represent times when the hippocampus efficiently communicates with the medial prefrontal cortex.

Acknowledgments

We thank Karola Käfer and Federico Stella for comments on an earlier version of the manuscript. This work was supported by the European Research Council (281511).

Author Contributions

H.X. and J.C. designed the experiments. H.X., P.B conducted the experiments and H.X. and J.O. analyzed the data. H.X. J.O. and J.C. wrote the manuscript.

Declaration of Interests

The authors declare no competing interests.

References

Ainge, J.A., Tamosiunaite, M., Wörgötter, F., and Dudchenko, P.A. (2012). Hippocampal place cells encode intended destination, and not a discriminative stimulus, in a conditional T-maze task. Hippocampus *22*, 534–543.

Allen, K., Rawlins, J.N.P., Bannerman, D.M., and Csicsvari, J. (2012). Hippocampal Place Cells Can Encode Multiple Trial-Dependent Features through Rate Remapping. J. Neurosci. *32*, 14752–14766.

Ambrose, R.E., Pfeiffer, B.E., and Foster, D.J. (2016). Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. Neuron *91*, 1124–1136.

Battaglia, F.P., Sutherland, G.R., and McNaughton, B.L. (2004). Local sensory cues and place cell directionality: additional evidence of prospective coding in the hippocampus. J. Neurosci. *24*, 4541–4550.

Csicsvari, J., Hirase, H., Czurko, A., Mamiya, A., and Buzsaki, G. (1999). Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving Rat. J.Neurosci. *19*, 274–287.

Csicsvari, J., O'Neill, J., Allen, K., and Senior, T. (2007). Place-selective firing contributes to the reverse-order reactivation of CA1 pyramidal cells during sharp waves in open-field exploration. Eur.J.Neurosci. *26*, 704–716.

Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal Replay of Extended Experience. Neuron *63*, 497–507.

Diba, K., and Buzsaki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. Nat.Neurosci. *10*, 1241–1242.

Dupret, D., O'Neill, J., Pleydell-Bouverie, B., and Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nat.Neurosci. *13*, 995–1002.

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron *23*, 209–226.

Ferbinteanu, J., and Shapiro, M.L. (2003). Prospective and retrospective memory coding in the hippocampus. Neuron *40*, 1227–1239.

Foster, D.J., and Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature *440*, 680-683.

Frank, L.M., Brown, E.N., and Wilson, M. (2000). Trajectory encoding in the hippocampus and entorhinal cortex. Neuron *27*, 169–178.

Grosmark, A.D., and Buzsáki, G. (2016). Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. Science *351*, 1440–1443.

Harris, K.D., Henze, D.A., Csicsvari, J., Hirase, H., and Buzsaki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J.Neurophysiol. *84*, 401–414.

Hasselmo, M.E. (1999). Neuromodulation: acetylcholine and memory consolidation. Trends Cogn Sci *3*, 351–359.

Hok, V., Lenck-Santini, P.P., Roux, S., Save, E., Muller, R.U., and Poucet, B. (2007). Goal-related activity in hippocampal place cells. J. Neurosci. *27*, 472–482.

Hollup, S.A., Molden, S., Donnett, J.G., Moser, M.B., and Moser, E.I. (2001). Accumulation of hippocampal place fields at the goal location in an annular watermaze task. J.Neurosci *21*, 1635–1644.

Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579.

Jadhav, S.P., Kemere, C., German, P.W., and Frank, L.M. (2012). Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory. Science *336*, 1454–1458.

Johnson, A., and Redish, A.D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. *27*, 12176–12189.

Karlsson, M.P., and Frank, L.M. (2009). Awake replay of remote experiences in the hippocampus. Nat.Neurosci. *12*, 913–918.

Lee, A.K., and Wilson, M.A. (2002). Memory of sequential experience in the hippocampus during slow wave sleep. Neuron *36*, 1183–1194.

Lee, I., Griffin, A.L., Zilli, E.A., Eichenbaum, H., and Hasselmo, M.E. (2006). Gradual Translocation of Spatial Correlates of Neuronal Firing in the Hippocampus toward Prospective Reward Locations. Neuron *51*, 639–650.

Leutgeb, S., Leutgeb, J.K., Treves, A., Moser, M.B., and Moser, E.I. (2004). Distinct ensemble codes in hippocampal areas CA3 and CA1. Science *305*, 1295–1298.

Mamad, O., Stumpp, L., McNamara, H.M., Ramakrishnan, C., Deisseroth, K., Reilly, R.B., and Tsanov, M. (2017). Place field assembly distribution encodes preferred locations. PLoS Biol. *15*, e2002365.

Markus, E.J., Qin, Y.L., Leonard, B., Skaggs, W.E., McNaughton, B.L., and Barnes, C.A. (1995). Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J.Neurosci. *15*, 7079–7094.

McNamara, C.G., Tejero-Cantero, Á., Trouche, S., Campo-Urriza, N., and Dupret, D. (2014). Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. Nat. Neurosci. *17*, 1658–1660.

Mehta M.R., McNaughton B.L. (1997) Expansion and Shift of Hippocampal Place Fields: Evidence for Synaptic Potentiation during Behavior. In: Bower J.M. (eds) Computational Neuroscience. Springer, pp. 741-745.

Mehta, M.R., Barnes, C.A., and McNaughton, B.L. (1997). Experience-dependent, asymmetric expansion of hippocampal place fields. Proc.Natl.Acad.Sci.U.S.A *94*, 8918–8921.

Mirenowicz, J., and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature *379*, 449–451.

Monaco, J., Rao, G., Roth, E., and Knierim, J.J. (2014). Attentive Scanning BehaviirAttentive Scanning Behavior Drives One-Trial Potentiation of Hippocampal Place Fields. Nat. Neurosci. *17*, 725–731.

Morris, R.G., Garrud, P., Rawlins, J.N., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. Nature *297*, 681–683.

Muller, R.U., and Kubie, J.L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J.Neurosci. *7*, 1951–1968.

Navratilova, Z., Hoang, L.T., Schwindel, C.D., Tatsuno, M., and McNaughton, B.L. (2012). Experience-dependent firing rate remapping generates directional selectivity in hippocampal place cells. Front. Neural Circuits *6*, doi: 10.3389/fncir.2012.00006.

Nitz, D.A., and McNaughton, B.L. (1999). Hippocampal EEG and unit activity responses to modulation of serotonergic median raphe neurons in the freely behaving rat. Learn.Mem. *6*, 153–167.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res *34*, 171–175.

O'Keefe, J., and Nadel, L. (1978). Hippocampus as a Cognitive Map, Clarindon.

Olton, D.S., and Samuelson, R.J. (1976). Remembrance of places passed: Spatial memory in rats. J. Exp. Psychol. Anim. Behav. Process. *2*, 97.

O'Neill, J., Senior, T., and Csicsvari, J. (2006). Place-selective firing of CA1 pyramidal cells during sharp wave/ripple network patterns in exploratory behavior. Neuron *49*, 143–155.

O'Neill, J., Boccara, C.N., Stella, F., Schoenenberger, P., and Csicsvari, J. (2017). Superficial layers of the medial entorhinal cortex replay independently of the hippocampus. Science *355*, 184–188.

Papale, A.E., Zielinski, M.C., Frank, L.M., Jadhav, S.P., and Redish, A.D. (2016). Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. Neuron *92*, 975–982.

Pepeu, G., and Giovannini, M.G. (2004). Changes in acetylcholine extracellular levels during cognitive processes. Learn. Mem. Cold Spring Harb. N *11*, 21–27.

Pfeiffer, B.E., and Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. Nature *497*, 74–79.

Rawlins, J.N., and Olton, D.S. (1982). The septo-hippocampal system and cognitive mapping. Behav. Brain Res. *5*, 331–358.

Redish, A.D. (2016). Vicarious trial and error. Nat. Rev. Neurosci. *17*, 147–159.

Sasaki, T., Piatti, V.C., Hwaun, E., Ahmadi, S., Lisman, J.E., Leutgeb, S., and Leutgeb, J.K. (2018). Dentate network activity is necessary for spatial working memory by supporting CA3 sharp-wave ripple generation and prospective firing of CA3 neurons. Nat. Neurosci. *21*, 258–269.

Singer, A.C., Carr, M.F., Karlsson, M.P., and Frank, L.M. (2013). Hippocampal SWR Activity Predicts Correct Decisions during the Initial Learning of an Alternation Task. Neuron *77*, 1163–1173.

Spellman, T., Rigotti, M., Ahmari, S.E., Fusi, S., Gogos, J.A., and Gordon, J.A. (2015). Hippocampal-prefrontal input supports spatial encoding in working memory. Nature *522*, 309–314.

Takeuchi, T., Duszkiewicz, A.J., Sonneborn, A., Spooner, P.A., Yamasaki, M., Watanabe, M., Smith, C.C., Fernández, G., Deisseroth, K., Greene, R.W., et al. (2016). Locus coeruleus and dopaminergic consolidation of everyday memory. Nature *537*, 357–362.

Wikenheiser, A.M., and Redish, A.D. (2015). Hippocampal theta sequences reflect current goals. Nat. Neurosci. *18*, 289–294.

Wood, E.R., Dudchenko, P.A., Robitsek, R.J., and Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron *27*, 623–633.

Yamamoto, J., and Tonegawa, S. (2017). Direct Medial Entorhinal Cortex Input to Hippocampal CA1 Is Crucial for Extended Quiet Awake Replay. Neuron *96*, 217-227.

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J. Neurophysiol. *79*, 1017–1044.

Main Figure Legends

Figure 1. Behavioral tasks on the radial eight-arm maze requiring reference or working memories or the combination of both

(A) Left panels: Illustration of the behavioral paradigm. In the combined task (Combined), all eight arms were open, three of which were baited with rewards. In the reference memory task (Reference), each arm was blocked within a trial, following a visit by the animal (grey bars in the center). In the working memory task (Working) arms not containing rewards were blocked throughout the recording (black bars in the center). Right traces: Behavioral tracks of the animals during different trials. Pre-probe trials (Pre) tested the memory of the goal-arm configuration from the previous day. During learning trials and in the subsequent post-probe trials (Post) two new goal arms (red dots) were selected, while the third one remained the same (blue dot). Animals were rested for 30 min after the probe and learning trials. (B) Learning curve of the animals showing the number of arms visited by the animal per trial (mean±SEM. (C-D) Comparison of reference (C) and working (D) memory errors during learning (mean±SEM). Combined n=16; Reference: n=8; Working n=8.

Figure 2. Gradual reorganization of place fields during learning in the combined task

(A) Firing fields of representative place cells during different trial blocks (6 trials per block) of the learning and probe sessions (Pre and Post). Numbers on the top right represent the maximum rate (Hz) of the cell. (B-C) Place field similarity between firing fields in learning trial blocks and in the post-probe session. The field representing all visited arms (B) or fields on the novel- and familiar-goal arms were compared (C). Correlation between learning block and place field similarity, all n=80, r=0.6885, P<0.0001; familiar-goal n=80, r=0.1660, P=0.1412; novel-goal n=80, r=0.5293, P<0.0001. (D) The proportion of cells exhibiting directional firing at different learning days for the combined task. Proportions of unidirectional and bidirectional cells were similar in day one P=0.7313 but significantly different in other days all P <0.01, t-test, n=16 recording days. Errorbars: SEM. See also Figure S1-S2.

Figure 3. In the combined task, place fields gradually shift towards reward locations during reward-bound passes of novel-goal arms

(A) Example firing fields generated from different learning trials for reward-bound and center-bound runs in familiar- and novel-goal arms. Two cells of each category are shown. Maps show the rate maps of cells on the goal-arm during different trials. Note that the shift towards the reward only occurred in reward-bound passes on novel-goal arms and that the cell in the top-left started to fire from the third trial only. (B) Distribution of place field shift scores comparing the place field peak positions in the first and last five trials. Shift scores measured the difference between the positions divided by the sum. Shift score is only significant for novel-goal arm reward-bound runs. Binomial tests, novel-goal: reward-bound, n=417, P<0.001, center-bound, n=432, P=0.879; familiar-goal: reward-bound, n=375, P=0.922, center-bound, n=359, P=0.630. (C) Distribution of place field peak positions as measured in the first and last five trials for familiar and novel arms. Note that significantly more cells fired near the goals in the last five trials than in the first five trials, but only in reward-bound, novel-goal arm runs. One-tailed

Kolmogorov-Smirnov-test (K-S test), novel-goal: reward-bound, n=417, P<0.001, center-bound, n=432, P=8766; familiar-goal: reward-bound, n=375, P=0. 0.7943, center-bound, n=386, P=0.5884. **P<0.001, ns, not significant. See also Figure S3.

Figure 4. Firing rate modulation of cells depends on past or future arm choices

(A-B) The firing rate modulation of place cells during reward-bound runs depends on the previous arm choice (A) while, in center-bound runs, it predicts the future choice (B), a relationship that develops during the course of learning. Place field maps illustrate the firing rates on trajectories involving the other two goal-arms. In the second half of learning cells fired more robustly on a novel-goal arm when the other segment of the path was also a novel-goal arm, as compared to a familiar-goal arm. The mean (±SEM) firing rate are shown for different spatial bins in the top two rows. Bottom histogram: mean (±SEM) speed at different spatial bins. Blue curves: passes including a familiar- and a novel-goal arms, red curves: passes across novel-goal arms. Note that rate differences were stronger in the second half of learning. (C) Rate modulation of cells on a novel-goal arm on reward-bound (left) and center-bound (right) passes depending on the past or future arm choices, respectively. The distribution of firing rate scores compare passes in which the previous or future arms were either a familiar-goal arm or another novel-goal arm. Rate scores were calculated by the difference between the rates between trajectory types (within a single arm) divided by their sum. A positive rate score reflects an increase in place field firing when the animal had run from a novel arm to another novel arm, as compared between a familiar and novel arm combination, while negative rate scores denote a stronger firing in passes that combine familiar- and novel-goal arms. The inset illustrates the meaning of the rate scores on the histograms. The thicker arrows indicate which arm combinations of a trajectory exhibit higher rates on the maze segment highlighted in grey for negative or positive scores. N=novel arm F=Familiar arm. Rate score distributions were calculated in the first half (grey) and second half (back) of the learning trial. Binomial tests, reward-bound: first half n=457, P=0.644, second half n=457, P<0.001, center-bound: first half n=376, P=0.986, second half n=376, P<0.001. (D) Same as (C), but calculated with only those cells that show a significant, speed compensated rate modulation (ANCOVA, p<0.05). Binomial tests, reward-bound: first half, n=182, P=0.515, second half, n=182, P<0.001, center-bound: first half, n=151, P=0.428, second half, n=151, P<0.001. See also Figure S4.

Figure 5 Maze trajectory reactivation during HSEs during learning

(A) Example showing place cell assembly activity during a pass covering movement from one goal-arm to the next in the combined task. Left traces: Grey traces show all the movement during all the learning trials while the black trace shows the pass for which the network activity is displayed on the right. Pink circles mark the location of the two HSEs. Top right raster plot: firing times of 24 place cells sorted according to their place field location along the trajectory.

Pink areas highlight the two HSEs, which are expanded in the left and the right boxes. Note that firing during the first HSE at the beginning of the pass (located at the reward, marked by a blue square) shows place cells that fire in the reverse order to that seen during traversal of the arm, from the reward to the center. In contrast, the HSE in the central area replays the activity of the cells in the next selected arm in forward temporal order, i.e., firing occurs in the same order as that of the arm. The bottom right curve shows the speed of the animal during the pass. (B-C) Examples illustrating how the replayed trajectories were calculated using two HSEs, one near a reward location (B) and the other emitted in the central area (C). Bayesian probability maps for both reward-bound and center-bound runs are shown for all three goal-arms. Probabilities are normalized for each time window by dividing by the sum. The map composed of reward-bound and center-bound maps of the selected arm is shown on the bottom right, while the reconstructed trajectory is shown on the bottom left (see STAR Methods). (D) Examples of replayed trajectories in recordings during the combined task that occurred near the reward location and the center area. Bayesian maps are overlaid with the reconstructed trajectory. HSEs at the reward represent center-bound passes of that arm (in reverse order), while HSEs in the center encode reward-bound passes of the next arm (forward order). The distance relative to the HSE location was displayed on the probability maps, i.e., zero marks the end of the arm for HSEs near reward locations and beginning of the arm for HSEs in the central area.

Figure 6 HSEs represent trajectories with significantly better replay scores than their shuffled equivalents and they tend to occur near SWRs

(A) The cumulative distribution of replay scores for the original events (black) and those of the shuffled events (grey) in which the place field relationship between cells was randomized by randomly rotating them to a different degree for each cell. Hundred random shuffled events were calculated for each original event. One-tailed K-S test, Combined: original n=1200, shuffle n=120000, $P<0.0000001$; Reference: original n=605, shuffle n=60500, $P<0.0000001$; Working: original n=691, shuffle n=69100, $P<0.0000001$; *** $P<0.0000001$. (B) Cross-correlation of SWR times with HSE peak times (mean±SEM).

Figure 7 Reactivated trajectories predicting the behavior of the animal

(A) Left: Example showing all the reactivated trajectories of HSEs that occurred on the central area during learning in a combined task session. Trajectories predicting the next arm choice of the animal (purple), the previous choice (green) and the rest of the events (black) are shown. Reactivated events are illustrated as a line showing the two extreme ends of the replayed trajectory. Therefore, the line length shows the length of the arm area covered by the reactivated trajectory. Different trajectories of the same arm are displayed by different relative angles from each other. Right: Proportion (mean ±SEM and individual session values) of trajectories predicting the future arm choice of the animal and those displaying its past arms for

HSEs in the center area. Comb: combined task, Ref: reference memory task, Work: working memory task. (B) Same as (A) but the majority of HSEs near reward location predict current and future arms. One-way ANOVA tests between future and past/current choice, Combined: n=16; Reference: n=8; Working: n=8, all P <0.0001 (C) Right panel illustrate reactivation HSEs representing reward- and center-bound passes in which the trajectory is encoded in the forward or the reverse order. Letters represent different segments of a trajectory. Proportion (mean ±SEM and individual session values) of HSEs in which forward and reverse replay was detected representing future or past/current trajectories of reward-bound vs. center-bound passes. Note that the combined and reference memory HSEs in the center area predicted the future arm choice of the animal in the form of forward replay of reward-bound trajectories. However, in the remaining cases (replay at the rewards and during the working memory task), the previous/current arm choices were predicted, representing center-bound trajectories but in the reverse order. Three-way ANOVA with Tukey Post hoc, (Top) Combined and Reference: Future-Forward-Reward-bound group is significantly different from other groups, all P<0.0000001; Working: Past-Reverse-Center-bound group was significantly different compared to other groups, all P<0.0000001; (Bottom) All three task: Current-Reverse-Center-bound group vs. other groups, all P<0.0000001. See also Figure S5-S7.

Figure 8

Reactivation content at correct and error passes

(A) The proportion (mean ±SEM and individual session values, Combined: n=16; Reference: n=8; Working: n=8) of HSEs in which forward and reverse replay was detected representing future or past/current trajectories of reward-bound vs. center-bound passes. Correct passes (top) and incorrect passes (bottom) are shown. For correct passes in the combined and reference memory tasks, the Future-Forward-Reward-bound group was higher than other groups, all P<0.0000001, while, in the working task, the Past-Reverse-Reward-bound group was the highest, all P<0.0000001; in error passes no significant differences between groups were seen all P>0.112067 (three-way ANOVA with Tukey Post hoc).

(B) The proportion of HSEs occurring at different relative locations in the central area of the maze during correct (top) and incorrect passes (bottom). Trajectory distance was measured relative to the entrance point to the central area and normalized by the length of the entire trajectory from the entrance to exit from the central area. Histograms for correct and error passes were significantly different (all P<0.031046, K-S Test). See also Figure S8.

STAR Methods

Contact for Reagent and Resource Sharing

Further information and requests for resources and datasets should be directed to and will be fulfilled by the Lead Contact, Jozsef Csicsvari (jozsef.csicsvari@ist.ac.at).

Experimental Model and Subject Details

Ten adult male rats (Long Evans, 300-500g, 3-6 months old) were used in this study. The animals were housed in a separate room under a 12/12 h light/dark cycle and were taken to the recording room each day prior to the experiments. They shared a cage with littermates before the surgery. All procedures involving experimental animals were carried out in accordance with Austrian animal law (Austrian federal Law for experiments with live animals) under a project license approved by the Austrian Federal Science Ministry.

Method Details

Animals and Surgery

Animals were implanted with 16 independently-movable wire tetrodes under deep anesthesia using isoflurane (0.5%–3%), oxygen (1–2 l/min), and an initial dose of buprenorphine (0.1 mg/kg). Tetrodes were attached to the 16-tetrode microdrive assemblies, enabling their independent movement. The tetrodes were constructed from 4 individual tungsten wires, 12 µm in diameter (H-Formvar insulation with Butyral bond coat, California Fine Wire, Grover Beach CA), twisted and then heated in order to bind them into a single bundle. The tips were then gold plated to reduce their impedance to 200 - 300 kΩ. During surgery, the craniotomy was centered above the CA1 at 2.5 mm (medial-lateral axis, ML) by -3.5mm (anterior-posterior, AP) and the electrodes were then implanted into deep layers of the neocortex. Two screws positioned above the cerebellum served as ground and reference electrodes. Six additional stainless-steel anchor screws were used in order to attach the microdrive assembly to the skull permanently. The electrodes and the microdrive apparatus were paraffin wax coated and daubed with dental acrylic to encase the electrode-microdrive assembly and anchor it to the screws in the skull. Following a recovery period of 7 days, the tetrodes were lowered in 50–200 µm steps each day into the CA1 region, over a further period of up to 7–14 days. Two 32-channel unity-gain preamplifier panels were used to reduce cable-movement artifacts. Wide-band (0.4 Hz to 9 kHz) recordings were taken, and the amplified local field potential and multiple-unit activity were continuously digitized at 24 kHz (Axona Ltd, St. Albans, UK).

Apparatus and training procedures

Following the recovery period (7 days), the animal was exposed to the radial eight-arm maze apparatus, over a period of 14 days. The radial eight-arm maze apparatus comprised of eight side arms (70cm long and 12cm wide) and an octagonal central area (80 cm high). Pneumatically controlled doors could be raised to prevent the entry of the animal to any of the side arms. Animals were put on food restriction (>85 % of initial weight with a 10 gram gain each week), once they had passed the seven day recovery period and exceeded their pre-surgery weight. Each animal was then trained to collect rewards on the eight-arm maze. Four animals were trained for the combined task and three additional animals performed the

working memory and reference memory tasks. The description of the tasks and the training for these procedures were as follows.

For the Combined working and reference memory task, in the beginning of the trial, animals stayed in the central area with all arms blocked by the side-doors. Food rewards (20mg food pellets) were placed in the three rewarded arms after which all side doors were simultaneously lowered. The trial ended when the animal collected all three rewards. Once the animal collected the last food reward, all the doors were raised except the one allowing the animal to return to the center area. After the animal returned to the central area, this remaining side door was raised and the next trial started after a one minute delay. Each experimental day started with the pre-probe session during which animals were tested over five trials with the reward-arm configuration of the previous day. After the pre-probe session, the animal was rested in the central area for 30min. After rest, the learning session started in which two of the reward locations were moved to arms that were not rewarded on the previous day (i.e., novel-goal arms) while one of the rewarded arms remained the same (i.e., familiar-goal arm). During the learning session, 30 trials were performed, which was sufficient for the animal to learn the new reward-arm configuration and perform >10 error-free trials. The learning session was followed by an additional 30 min rest period and a subsequent post-probe session when the newly-learned reward-arm configuration was tested again for an additional five trials.

To train the animals for this combined task (four animals), initially, for two training days, all eight side-arms contained food and the animal had to collect food on all eight arms before a trial ended and the side arms were rebaited. During this initial training, 10-15 of these eight-arm baited trials were performed. After this, for the next 4-5 training days, animals had to perform the three-arm baited task until the animal performed at least ten trials successfully without errors (maximum 50 trials per day). During training, not only the learning but also the probe sessions were performed. In the final training day it was required that the animal was able to perform ten trials without errors within the first 30 trials, otherwise, a further training day was added before recording could start.

As in the combined task, in both the working and reference memory tasks two of the goal-arms were switched daily and one remained the same. Also, the same sessions were used in the same order as in the combined task. For the working memory task, only three arms of the maze were opened and a trial ended once the animal collected the food from all three arms. For the reference memory tasks the procedure was identical to the combined task with the exception that once the animal returned from an arm, that arm was blocked from re-entering by raising the connecting door at the central area.

For both the working and reference memory tasks the same animals were used (three animals), but these were different from the one ones used for the combined task. Training procedure during the first two days was identical to that of the combined task, i.e., sessions in which all eight arms were baited. After this, only one training day was needed to familiarize the animal with the working memory task, and subsequently, four recording days were performed. After the four recording days with the working memory task, animals were trained to the reference memory task. Training procedure was like in the combined task; only within each trial, the visited arms were blocked after the animal returned from them.

Because the training order in the working and reference memory tasks may have influenced our results, we performed additional experiments in a further three animals which were trained to perform the reference memory first followed by the working memory task. Recording and training protocols were identical to those described above for the original experiments.

Spike sorting

The spike detection in the local field potential and sorting was performed as previously described (Csicsvari et al., 1999). Action potentials were extracted by first computing power in the 800-9000 Hz range within a sliding window (12.8 ms). Action potentials with a power of >5 SD from the baseline mean were selected and spike features were then extracted by using principal components analyses. The detected action potentials were then segregated into putative multiple single units by using automatic clustering software (http://klustakwik.sourceforge.net/) (Harris et al., 2000). These clusters were then manually refined by a graphical cluster cutting program. Only units with clear refractory periods in their autocorrelation and well-defined cluster boundaries were used for further analysis. We further confirmed the quality of cluster separation by calculating the Mahalanobis distance (Harris et al., 2000) between each pair of clusters. Periods of waking spatial exploration, immobility, and sleep were clustered together and the stability of the isolated clusters was examined by visual inspection of the extracted features of the clusters over time. Pyramidal cells and interneurons in the CA1 region were discriminated by their autocorrelations, firing rate, and waveforms, as previously described (Csicsvari et al., 1999). In this way, we were able to identify the activity of 2430 CA1 pyramidal units.

Quantification and Statistical Analysis

Generation of rate maps

Two-dimensional place-rate maps were calculated by first, subdividing the environment on a grid, containing 2x2cm bins. A 2x2cm square was overlaid on this grid, for each spike of a given cell, centered on the position of the animal when the spike was emitted. Each bin was then incremented by the degree to which this square overlapped with it. The same procedure was then performed with the animal position tracking data to produce a map of occupancy. Only spikes and occupancy periods that occurred during movement (5cm/s) were considered. Occupancy and spike matrices were then smoothed with a Gaussian kernel (SD = 3 bins), and divided to produce place-rate maps. To test for the rate modulation of trial type, place field shift and replay using Bayesian decoding, one-dimensional (1D) place maps were calculated using a similar procedure as above for each arm of the maze. Moreover, place maps were calculated separately for reward-bound and center-bound passes for the 1D maps. For the Bayesian reconstruction analysis, 70bins, i.e., 1cm bins were used for each arm.

ANCOVA analysis

The ANCOVA analysis was performed using spike firing rates established on the tested arm and movement direction. The tested arm was divided into five spatial bins along the path, in which rate and speed were calculated on each trial in which the animal visited the arms with the appropriate previous or next arm visits. Both rates at different spatial bins and speed were

calculated separately for cases when the animal was visiting the two other arms before or next. Only units that had larger than 1 Hz mean rate on the tested arm were included in the analysis. We then used an ANCOVA (using R software package) in order to ask if rate varied with trial stage given changes in the speed. We used spatial bins and arms (i.e., the two alternative arms to visit next or before) as factors and speed as a linear variable. We accepted cases as significant in which rate of the arms or arm vs. spatial bin interaction were significant, independent of speed (Allen et al., 2012)

High Synchrony Events (HSE)

High synchrony periods were detected using the multi-unit activity of clustered spikes, similar to that previously described (Davidson et al., 2009). The combined activity of pyramidal cells was binned into 1ms bins and smoothed with a Gaussian kernel (SD = 15 ms), to produce a curve representing the synchronous spiking rate over time. HSE detection began when this curve passed 3 SD above the mean. The envelope of the HSE was extended until the curve again reached the mean rate, either side of this crossing. The peak of the HSE represented the peak of the curve within this envelope. Events containing fewer than five spikes, four cells or with less than 10% of the population of neurons active were rejected, as were events shorter than 75ms or greater than 750 ms. The beginning of the HSE was then adjusted to the time of the first spike. The HSE was then subdivided into 20ms windows, with a 10 ms overlap, until the last window containing a spike was reached. Following this procedure, all events with less than five windows were rejected from further analysis.

Replay during HSEs

Each detected HSE was subdivided into 20ms time windows and population vectors representing the spike counts of different pyramidal cells in a given time window were calculated. Then we used the Bayesian place prediction (Zhang et al., 1998) method to calculate a probability of different positions for each population vector of the event. Probability maps were calculated for each arm, both for reward-bound and center-bound trials. The formula below gave the probability that a given population vector represented a given place:

$$P(x|\mathbf{n}) = P(\mathbf{n}|x)P(x)/P(\mathbf{n}).$$

P(x) represents the probability that the animal is at a given location considering the exploration session was set to a uniform distribution not to bias our analysis by any place preference of the animal (Zhang et al., 1998). P(n|x) represents the conditional probability that a given spike count occurs at a location. This was estimated using the firing rates of the place-rate maps, assuming that the number of spikes follows a Poisson distribution. P(n), the normalizing constant, was used to ensure that P(x|n) summed up to one. As for the calculated place rate maps (see above), probability maps contained 70 spatial bins as well.

We then ran an optimization procedure to see which arm produces the maximum summed probabilities across the probability maps representing different time windows of the HSE. The procedure tried out different jump combinations along the consecutive time windows. However, it allowed a jump of up to 20 spatial bins across probability maps of neighboring time windows assuming that replayed trajectories can jump a maximum of 20 spatial bins in each

step. Moreover, jumps were allowed either in the forward or reverse order relative to the movement direction of the animal to test for both forward and reverse replays. For the replayed trajectory we selected the arm that yielded the maximum summed probabilities across all HSE time windows and the positions were extracted from the probability map positions that yielded the maximum summed probabilities. However, time windows with zero probabilities were excluded from the trajectories. We only considered trajectories with minimum four time windows and only those that exhibited sufficient movement by covering at least 20 bins and had an average speed larger than two bins per time window.

A trajectory was determined as reward-bound when the trajectory from the beginning to the end showed an average shift towards the end of the arm and center-bound when the opposite was observed. The model selection determined whether forward or reverse replay was selected, that is when probabilities moved across time windows according to the appropriate maps. For example, a reward-bound shift using reward-bound map would reflect forward replay, whereas the same movement established using a center-bound map would result in a reverse replay.

To check whether the trajectories we detected can be generated from the sequential firing of place cells with random spatial representations, we compared the summed probability of our procedure against shuffled data in which place fields of the cells were rotated on each arm for reward- and center-bound trials separately. For each cell, the place fields constructed were randomly shifted along the arm independently, with bins extending off the end of the arm wrapped to the start. Next, we established the maximized summed probabilities by using the same procedure as above with these fields, over 100 different shuffled events. This method preserved spike timing and firing rate statistics while disrupting the spatial coding of the cells used in the reconstruction. We then established a distribution of replay scores (summed probabilities divided by the number of time windows) across 100 shuffles. From this shuffled data we established a further score (replay score-z) showing the distance of replay score from the mean of the shuffled distribution by subtracting the mean of the shuffled data from the replay score and normalizing by its standard deviation (Grosmark and Buzsáki, 2016).

SWR detection

SWR detection was performed as previously described (Csicsvari et al., 1999). Local field potentials were band-pass filtered (150–250 Hz), and a reference signal (from a channel that did not contain ripple oscillations) was subtracted to eliminate common-mode noise (such as muscle artifacts). The power (root mean square) of the filtered signal was calculated for each electrode and summed across electrodes designated as being in the CA1 pyramidal cell layer. The SWR detection threshold (7 SD above baseline) was always set in the first available sleep session, and the same threshold used throughout.

Cross-correlation analysis was performed to detect temporal alignment between the detected SWRs and CA1 HSEs. For the SWR analysis, HSE times were taken as a reference and the number of detected SWRs was calculated in 25ms windows, centered on the peak of the HSE and extending to ± 500 ms. The cross-correlogram was normalized by the total number of detected HSEs.

Data and Software Availability

Data used in this study will be made available upon request by contacting the lead contact, Jozsef Csicsvari (jozsef.csicsvari@ist.ac.at).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Long Evans Rats	Janvier, France	RjOrl:LE
Software and Algorithms		
KlustaKwik	Harris et al., 2000	https://github.com/klusta-team/klustakwik/
Python	https://www.python.org	Python
R	https://www.r-project.org/	R
Other		
12um tungsten wires	California Fine Wire	M294520
Headstage amplifier	Axona, St. Albans, UK	http://www.axona.com/

Figure 1

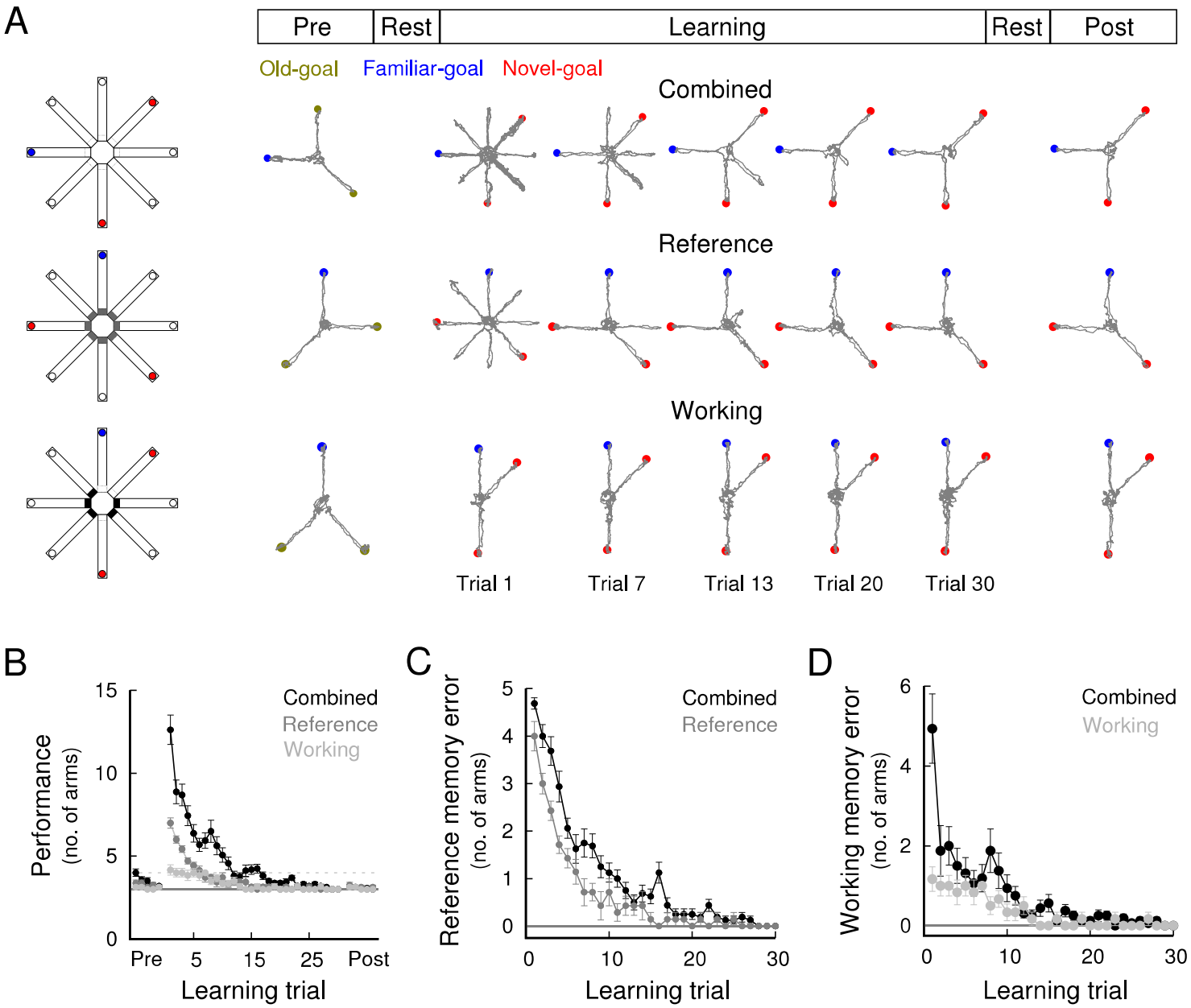

A
Pre
Rest
Learning
Rest
Post
Old-goal
Familiar-goal
Novel-goal
Combined
Reference
Working
Trial 1
Trial 7
Trial 13
Trial 20
Trial 30
B
Performance
(no. of arms)
Combined
Reference
Working
Pre
5
15
25
Post
Learning trial
C
Reference memory error
(no. of arms)
Combined
Reference
Learning trial
D
Working memory error
(no. of arms)
Combined
Working
Learning trial

Figure 2

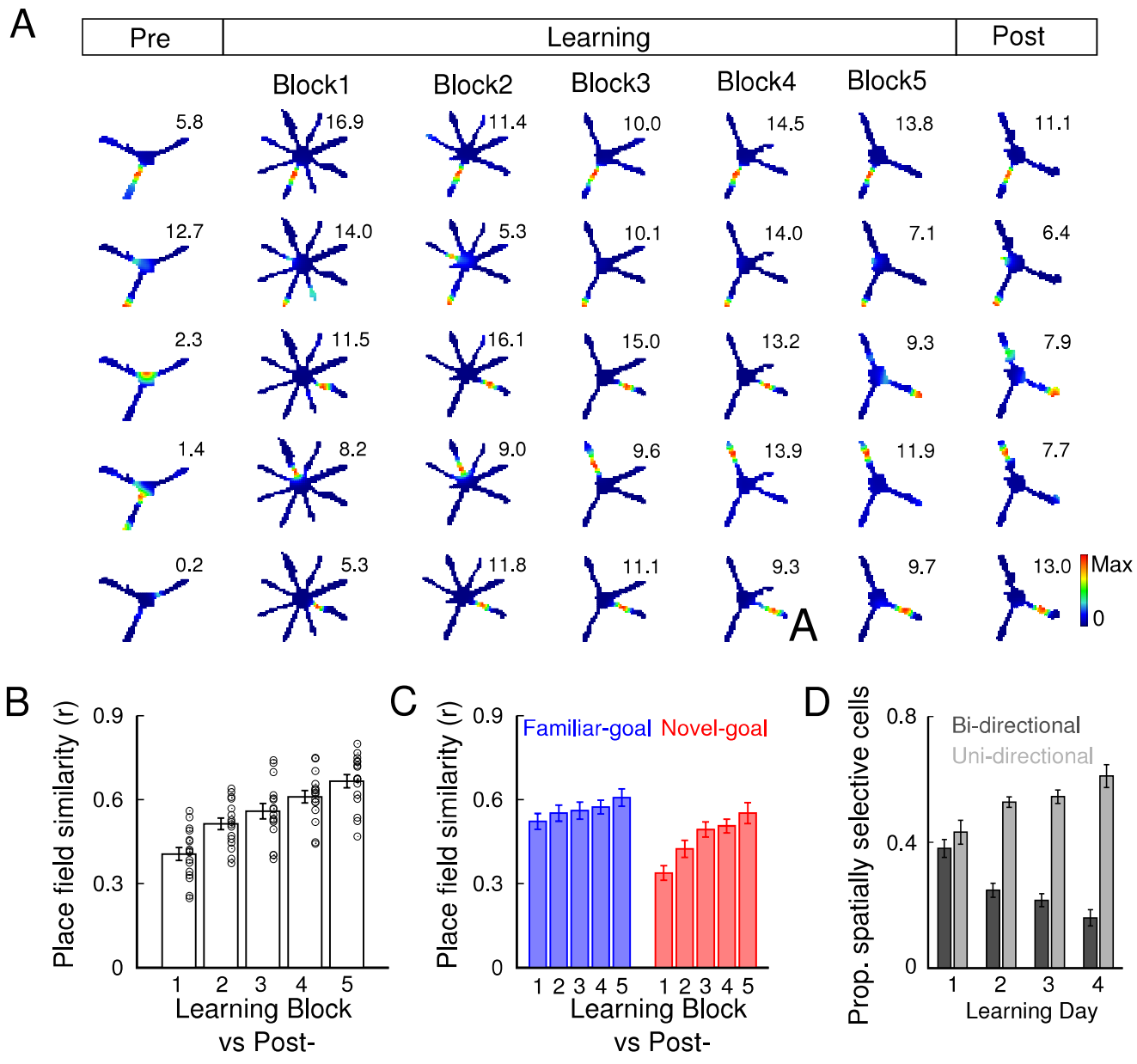

A
Pre
Learning
Post
Block1
Block2
Block3
Block4
Block5
5.8
16.9
11.4
10.0
14.5
13.8
11.1
12.7
14.0
5.3
10.1
14.0
7.1
6.4
2.3
11.5
16.1
15.0
13.2
9.3
7.9
1.4
8.2
9.0
9.6
13.9
11.9
7.7
0.2
5.3
11.8
11.1
9.3
9.7
13.0
Max
0
A
B
Place field similarity (r)
0.9
0.6
0.3
0
1 2 3 4 5
Learning Block
vs Post-
C
Place field similarity (r)
Familiar-goal
Novel-goal
1 2 3 4 5 1 2 3 4 5
Learning Block
vs Post-
D
Prop. spatially selective cells
0.8
0.4
0
Bi-directional
Uni-directional
1 2 3 4
Learning Day

Figure 3

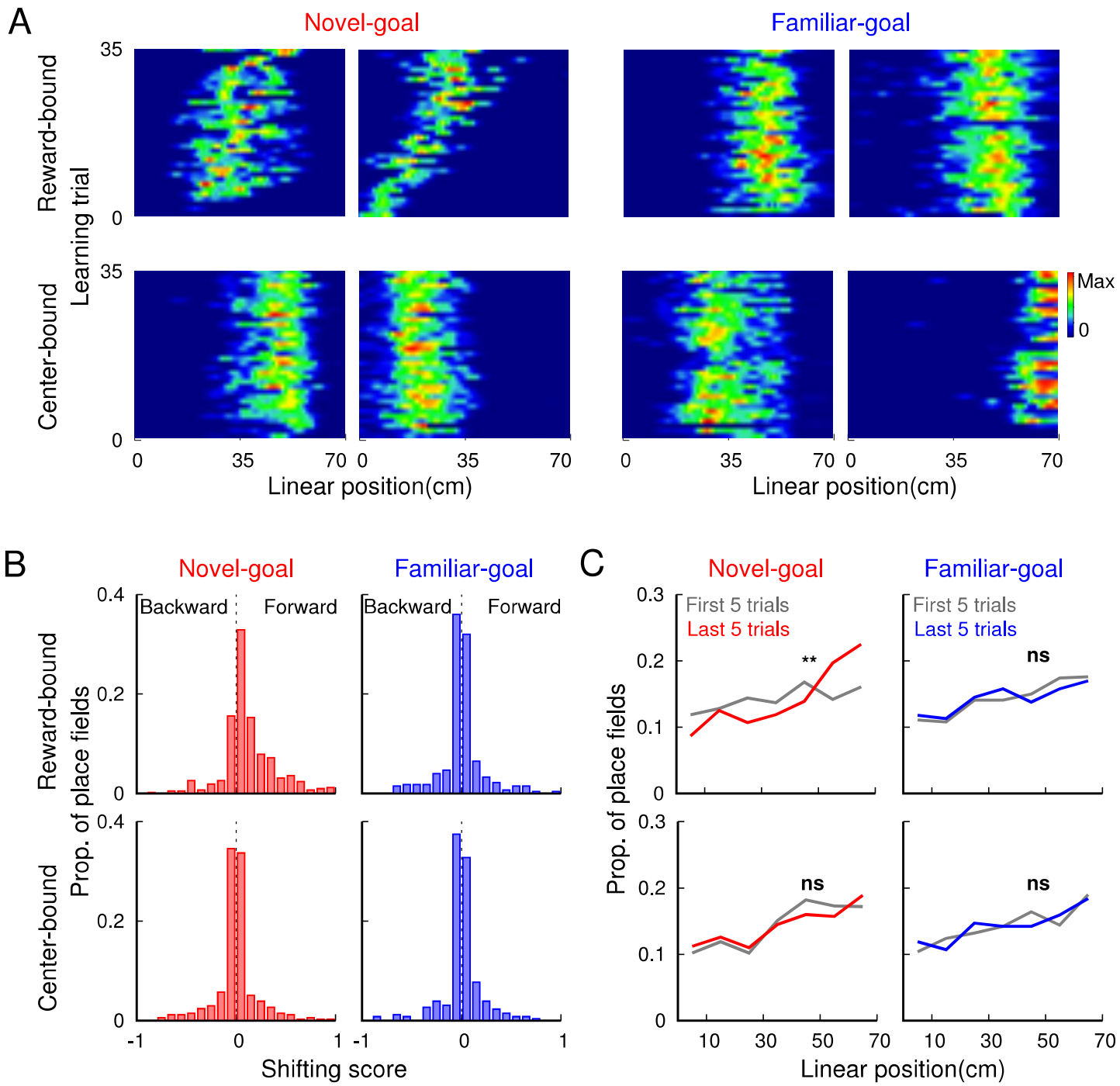

A
Novel-goal
Familiar-goal
Reward-bound
Center-bound
Learning trial
35
0
Max
0
0
35
70
Linear position(cm)
B
Novel-goal
Familiar-goal
Backward
Forward
Prop. of place fields
0.4
0.2
-1
1
Shifting score
C
First 5 trials
Last 5 trials
**
ns
0.3
0.1
10
30
50
70
Linear position(cm)

Figure 4

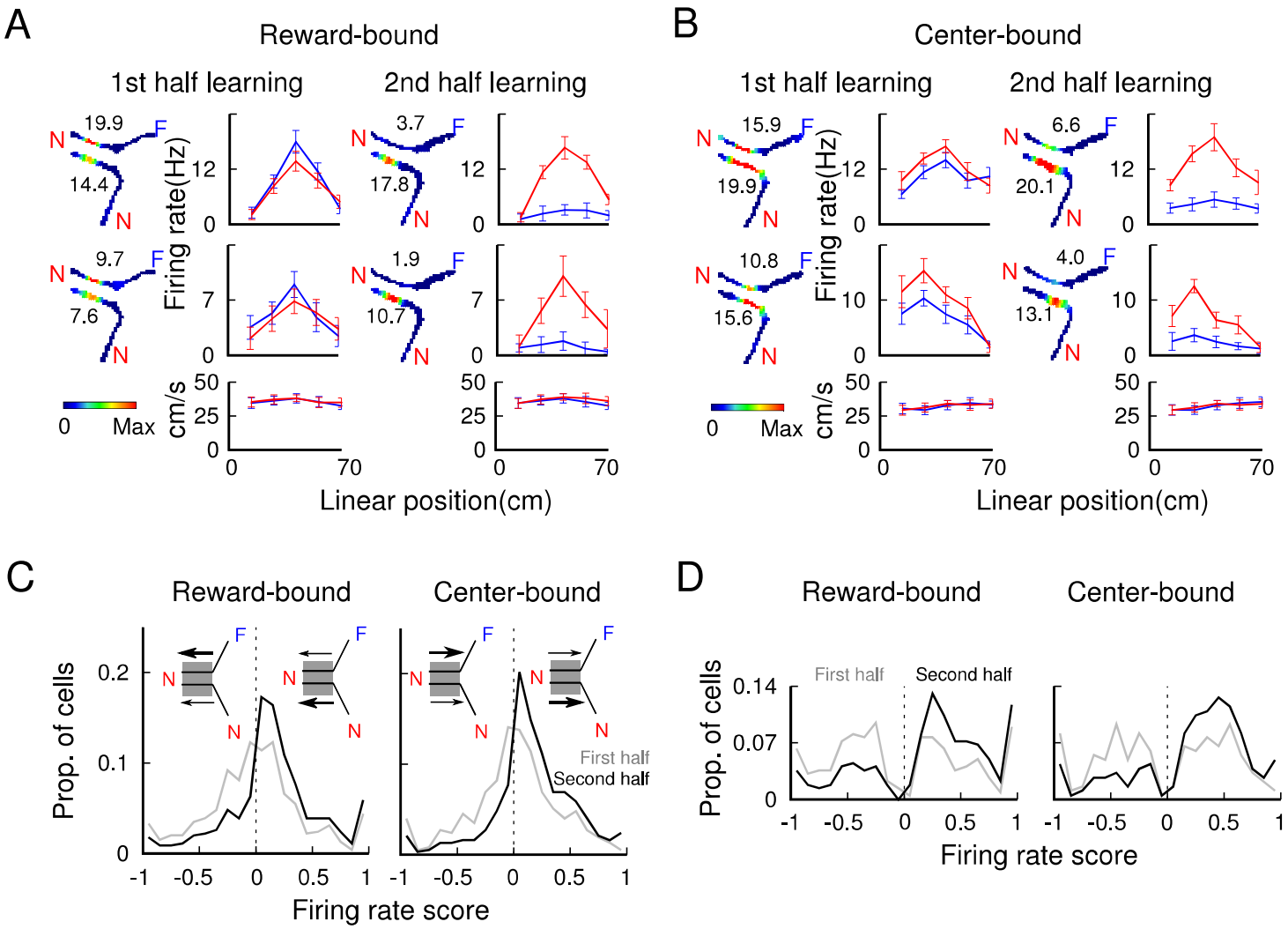

A
Reward-bound
1st half learning
2nd half learning
19.9
14.4
3.7
17.8
9.7
7.6
1.9
10.7
N
F
Firing rate(Hz)
12
7
0
0
Max
cm/s
50
25
0
70
Linear position(cm)
B
Center-bound
1st half learning
2nd half learning
15.9
19.9
6.6
20.1
10.8
15.6
4.0
13.1
Firing rate(Hz)
12
10
cm/s
Linear position(cm)
C
Reward-bound
Center-bound
Prop. of cells
0.2
0.1
First half
Second half
-1
-0.5
0.5
1
Firing rate score
D
Reward-bound
Center-bound
Prop. of cells
0.14
0.07
First half
Second half
Firing rate score

Figure 5

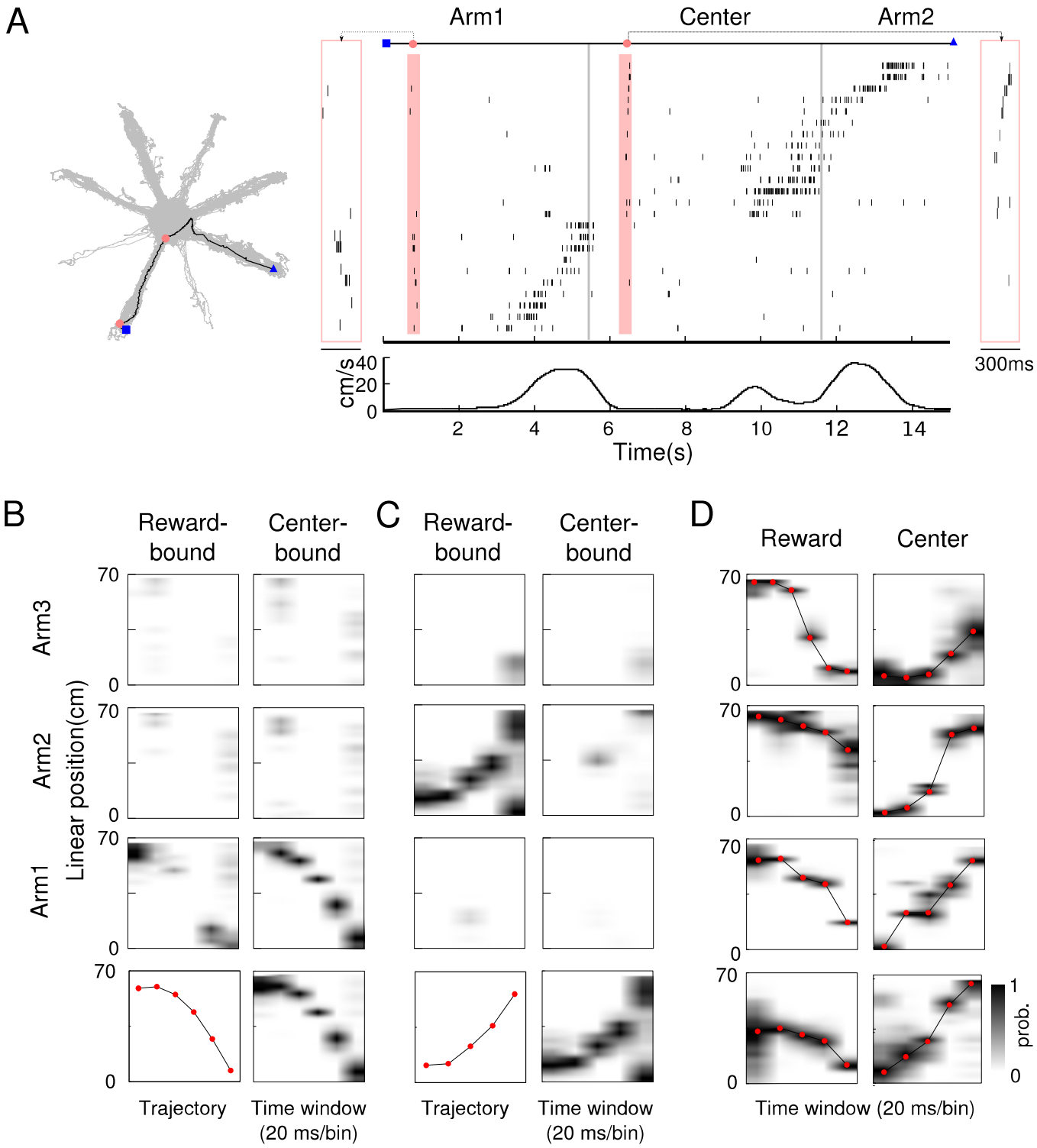

A
Arm1
Center
Arm2
300ms
cm/s
40
20
0
2
4
6
8
10
12
14
Time(s)
B
Reward-bound
Center-bound
C
Reward-bound
Center-bound
D
Reward
Center
Arm3
Arm2
Arm1
Linear position(cm)
70
0
Trajectory
Time window (20 ms/bin)
Trajectory
Time window (20 ms/bin)
Time window (20 ms/bin)
1
prob.
0

Figure 6

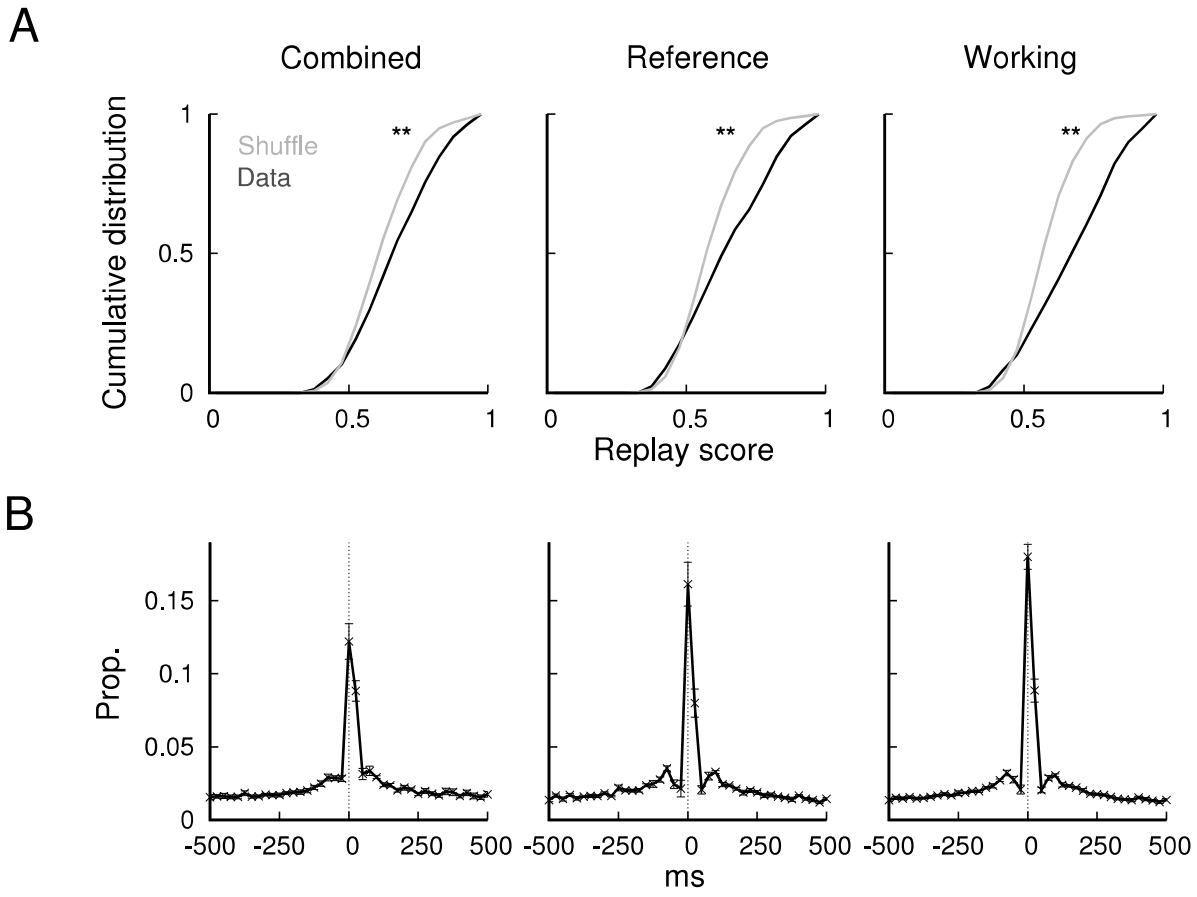

A
Combined
Reference
Working
Cumulative distribution
Shuffle
Data
**
**
**
1
0.5
0
0
0.5
1
Replay score
B
Prop.
0.15
0.1
0.05
0
-500
-250
0
250
500
ms

Figure 7

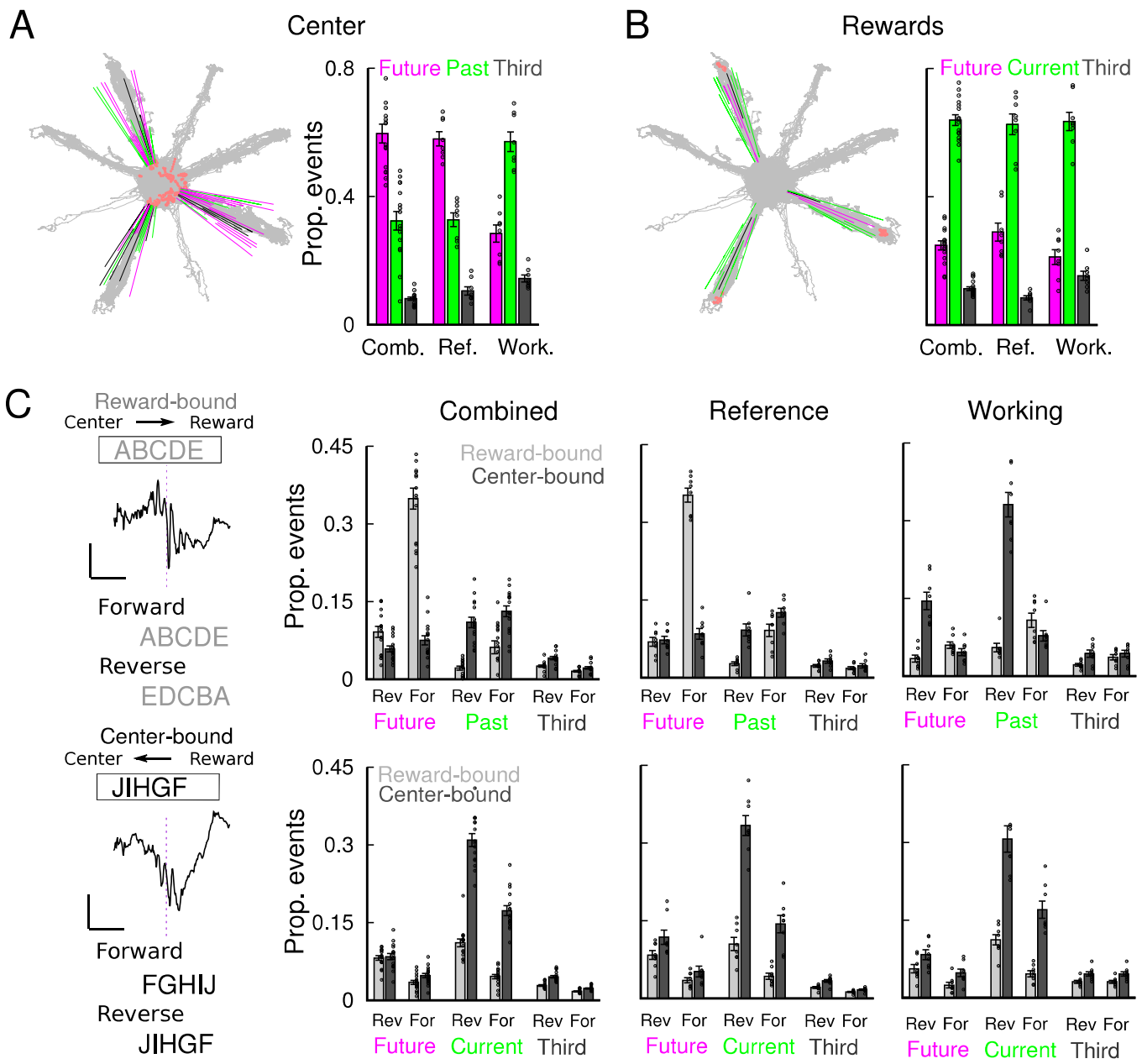
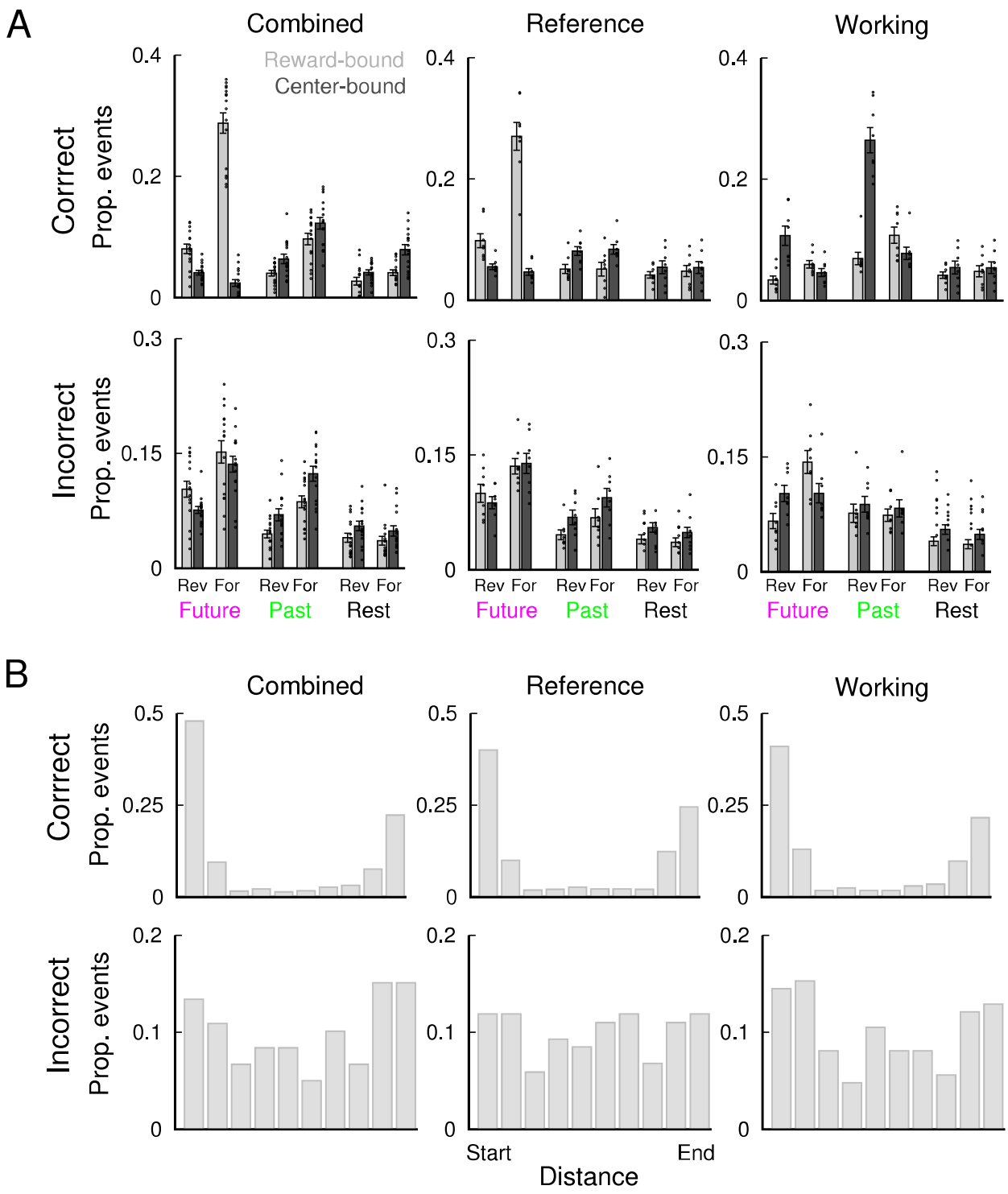

A
Center
Future Past Third
0.8
0.4
0
Prop. events
Comb. Ref. Work.
B
Rewards
Future Current Third
Comb. Ref. Work.
C
Reward-bound
Center → Reward
ABCDE
Forward
ABCDE
Reverse
EDCBA
Center-bound
Center ← Reward
JIHGF
Forward
FGHIJ
Reverse
JIHGF
Combined
Reference
Working
Reward-bound
Center-bound
0.45
0.3
0.15
0
Prop. events
Rev For Rev For Rev For
Future Past Third
Future Current Third

Figure 8


A
Combined
Reference
Working
Reward-bound
Center-bound
Corrrect
Prop. events
Incorrect
Prop. events
0.4
0.2
0
0.3
0.15
Rev For
Future
Past
Rest
B
Combined
Reference
Working
Corrrect
Prop. events
Incorrect
Prop. events
0.5
0.25
0.2
0.1
Start
End
Distance

Supplementary Figures

Assembly responses of hippocampal CA1 place cells predict learned behavior in goal-directed spatial tasks on the radial eight-arm maze

Haibing Xu, Peter Baracskay, Joseph O'Neill, Jozsef Csicsvari

Institute of Science and Technology Austria (IST Austria)

Am Campus 1, A – 3400 Klosterneuburg, Austria

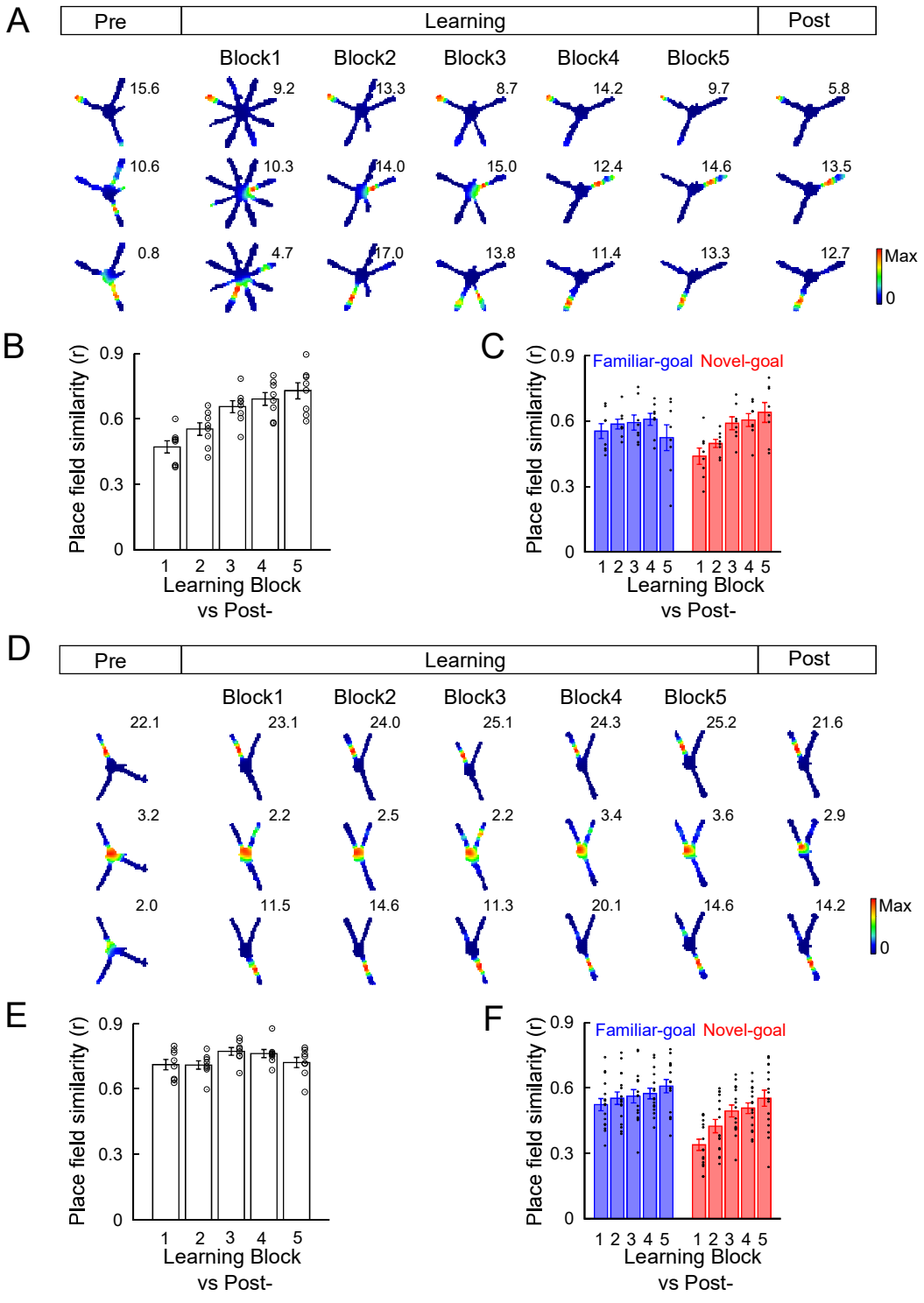

A
Pre
Learning
Post
Block1
Block2
Block3
Block4
Block5
15.6
9.2
13.3
8.7
14.2
9.7
5.8
10.6
10.3
14.0
15.0
12.4
14.6
13.5
0.8
4.7
17.0
13.8
11.4
13.3
12.7
Max
0
B
Place field similarity (r)
0.9
0.6
0.3
0
1 2 3 4 5
Learning Block
vs Post-
C
Familiar-goal Novel-goal
Place field similarity (r)
0.9
0.6
0.3
0
1 2 3 4 5 1 2 3 4 5
Learning Block
vs Post-
D
Pre
Learning
Post
Block1
Block2
Block3
Block4
Block5
22.1
23.1
24.0
25.1
24.3
25.2
21.6
3.2
2.2
2.5
2.2
3.4
3.6
2.9
2.0
11.5
14.6
11.3
20.1
14.6
14.2
Max
0
E
Place field similarity (r)
0.9
0.6
0.3
0
1 2 3 4 5
Learning Block
vs Post-
F
Familiar-goal Novel-goal
Place field similarity (r)
0.9
0.6
0.3
0
1 2 3 4 5 1 2 3 4 5
Learning Block
vs Post-

Figure S1. Place fields reorganize gradually during the reference memory task but not the working memory task, Related to Figure 2

(A,D) Firing fields of representative place cells during different learning trial blocks and probe sessions for reference memory (A) and working memory (D) tasks. Numbers represent the maximum firing rate (Hz) of the cell. (B-C, E-F) Place field similarity between firing fields in learning trial blocks and in the post-probe session for reference (B) and working (E) memory tasks comparing the entire firing field. Reference: n=40, r= 0.6994, P<0.0001; Working: n=40, r=0.1671, P=0.3028. Same shown for firing fields on novel- and familiar-goal arms for reference (C) and working (F) memory tasks. Reference: familiar-goal n=40, r=0.1420, P=0.9256; novel-goal n=40, r=0.5630, P<0.0001; Working: familiar-goal n=40, r=0.1834, P=0.1438; novel goal n=40, r=0.1238, P=0.8625.

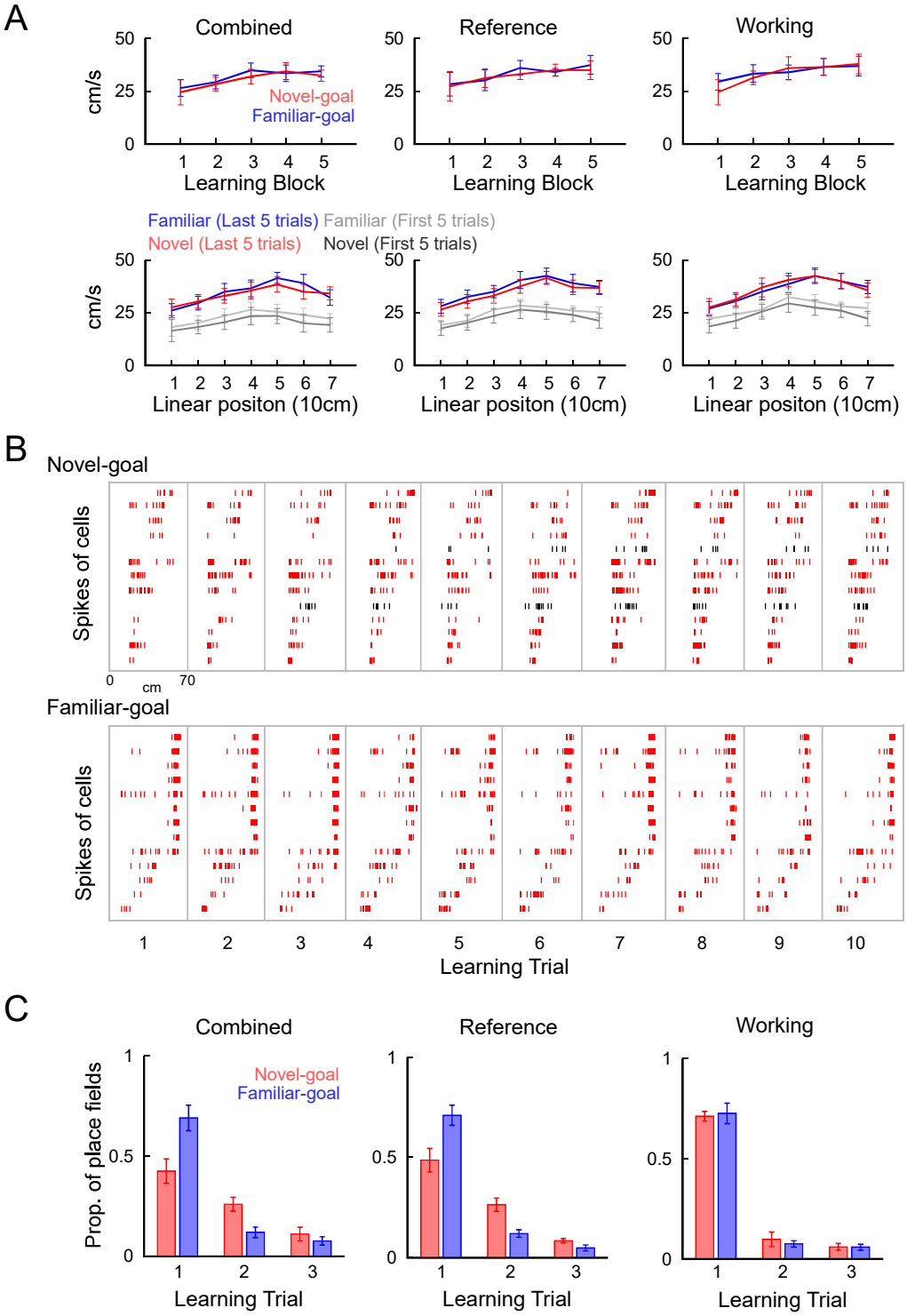

A
Combined
Reference
Working
cm/s
50
25
0
Novel-goal
Familiar-goal
1 2 3 4 5
Learning Block
Familiar (Last 5 trials)
Familiar (First 5 trials)
Novel (Last 5 trials)
Novel (First 5 trials)
1 2 3 4 5 6 7
Linear positon (10cm)
B
Novel-goal
Spikes of cells
0
cm
70
Familiar-goal
1 2 3 4 5 6 7 8 9 10
Learning Trial
C
Combined
Reference
Working
Prop. of place fields
1
0.5
0
Novel-goal
Familiar-goal
1 2 3
Learning Trial

Figure S2. Reorganization of place fields during learning, Related to Figure 2

(A) Mean (±SEM, Combined: n=16; Reference: n=8; Working: n=8) speed of the animal at different trial blocks (top) and at different locations on the arm comparing speed distributions in the beginning and the end of the trials (bottom). Speed values were similar on the familiar- and novel-goal arms but they increased slightly in later trials (Familiar. vs. Novel, all $P>0.4197$, learning blocks, all $P<0.01$ two-way ANOVA). Moreover, the speed profile of the animal as running on an arm was similar when novel and familiar arms were compared either in the beginning or the end of learning (Familiar vs. Novel, all $P>0.502$ three-way ANOVA). (B) Example showing the firing patterns of different place cells on a novel-goal arm (top) and familiar goal arm (bottom). The raster plot shows the arm locations at which the displayed cells fired at different trials. Each line represents a different cell. Note that the majority of cells fired from the beginning (red) but some started to fire at later trials (black). (C) The proportion of cells that started to fire in the first three trials in the combined working and reference memory tasks Combined: Novel n=516, Familiar n=349; Reference: Novel n=253, Familiar n=181; Working: Novel n=197, Familiar n=157. In the combined and reference memory tasks, a larger proportion of place fields emerged in the first trial in the familiar arm while the opposite was observed in the second trial (all $P<0.0102$ two-way ANOVA, Tukey Post hoc). In the third trials for all tasks as well as trial 1-2 of the working memory task there were no differences ($P > 0.5937$). Note that only 50-55% of the cells fired in the first trial in the combined and reference memory tasks on the novel-goal arm, while 75-80% fired in the familiar-goal arms and in the working memory task.

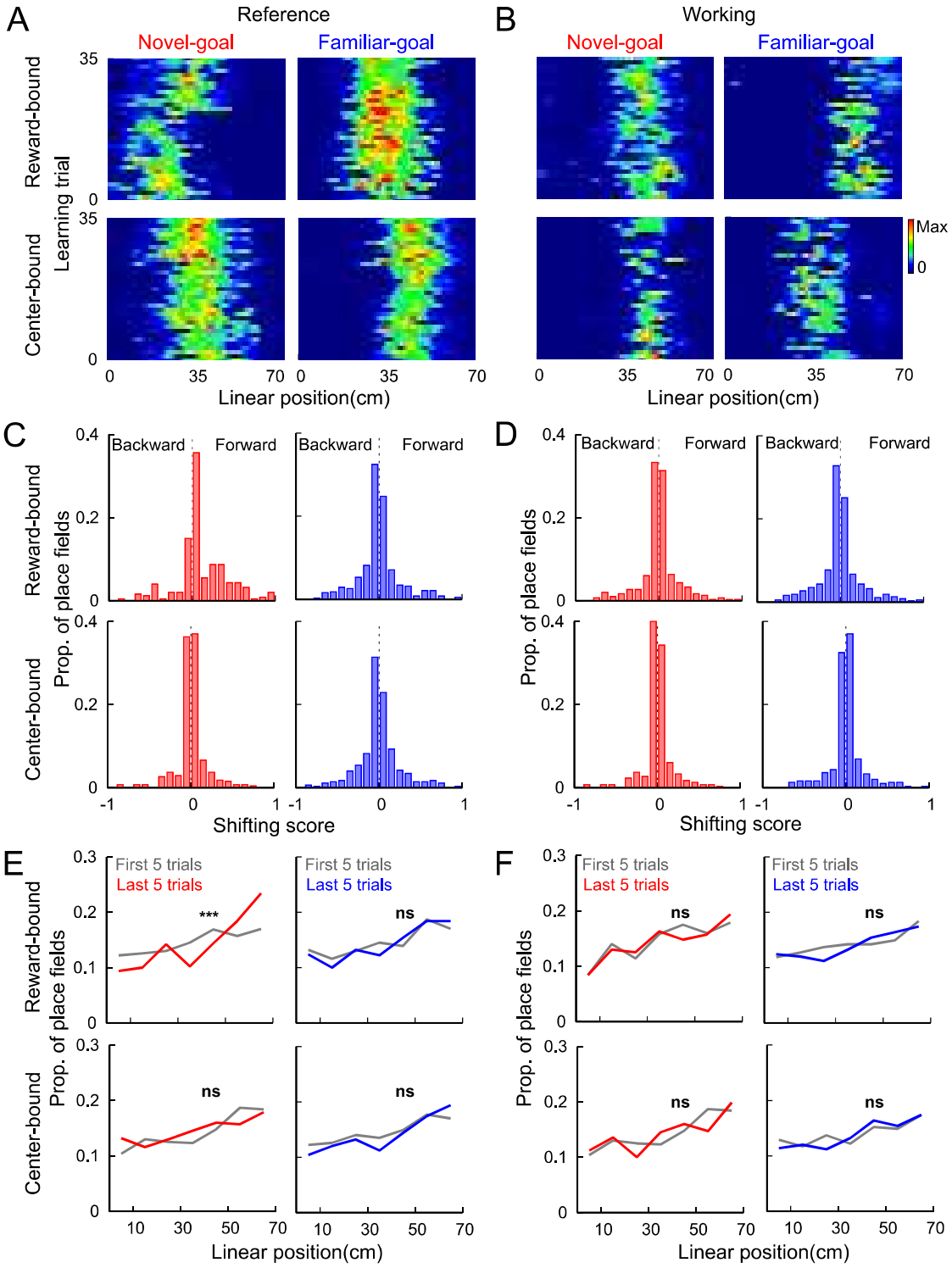

A
Reference
Novel-goal
Familiar-goal
Reward-bound
Center-bound
Learning trial
35
0
Linear position(cm)
B
Working
Max
C
D
Backward
Forward
Prop. of place fields
Shifting score
E
F
First 5 trials
Last 5 trials
***
ns

Figure S3. Place field shifts towards the reward locations is seen in the reference memory task but not in the working memory task, Related to Figure 3

(A-B) Example firing fields at different learning trials for reward-bound and center-bound trials in familiar- and novel-goal arms. One cell of each category is shown for the reference (A) and the working (B) memory tasks. The rate maps show the firing rate of the cells on the goal-arm at different trials. Note that shifts towards rewards only occurred during reward-bound passes on the novel-goal arms in the reference memory task. (C-D) Distribution of place field shift scores comparing the place field peak positions in the first and last five trials for reference (C) and working (D) memory sessions. Shift scores measured the difference between the positions divided by the sum. Binomial tests, Reference: novel-goal, reward-bound, n=253, P<0.001, center-bound n=275, P=0.517; familiar-goal, reward-bound n=353, P=0.819; center-bound, n=379, P=0.907. Working: novel-goal, reward-bound, n=312, P=0.830; center-bound, n=328, P=0.717; familiar-goal, reward-bound, n=382, P=0.678; center-bound, n=305, P=0.522. (E-F) Distribution of place field peak positions as measured in the first and last five trials for reference (E) and working (F) memory session. Note that more cells fired near the goal in the last five trials than in the first five trials but only in the reward-bound novel-goal runs during the reference memory task. One-tailed K-S tests for both tasks.
Reference: novel-goal, reward-bound n=253, P<0.0001, center-bound, n=275, P=0.7256; familiar-goal, reward-bound, n=353, P=0. 4180, center-bound n=379, P=0.2239. Working: novel-goal, reward-bound n=312, P=0.8961, center-bound n=328, P=0.7905; familiar-goal, reward-bound n=382, P=0.3077, center-bound n=305, P=0.7972. ***P<0.0001, ns, not significant.

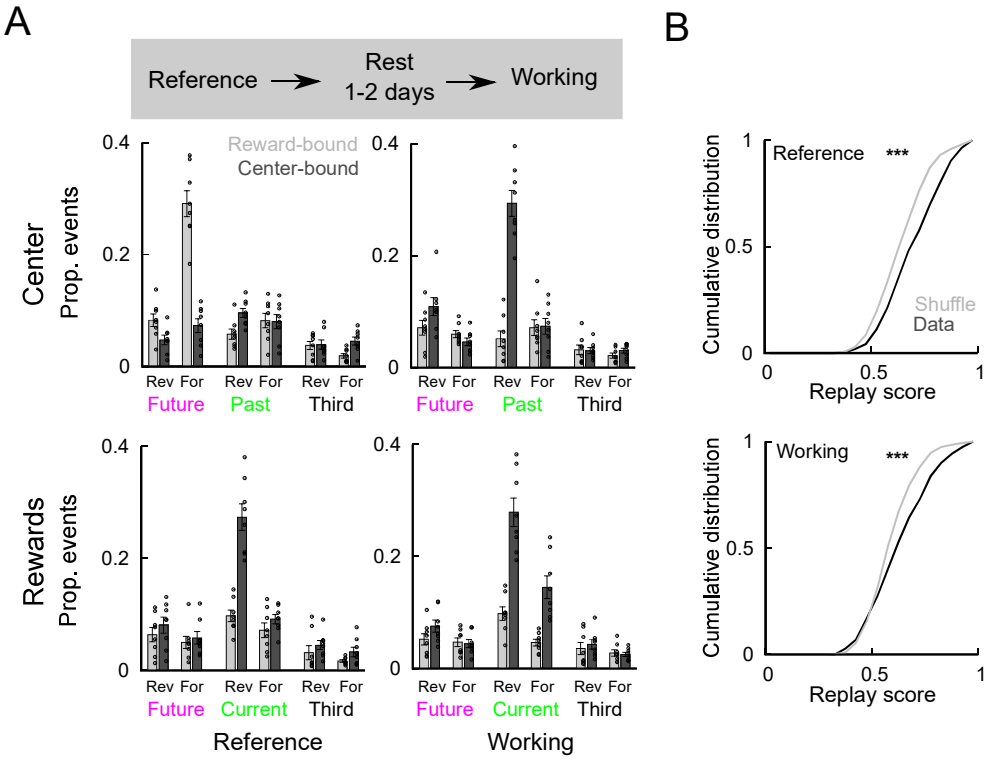


Figure S4. Previous or future arm choice-dependent firing rate modulation of cells on the reference and working memory tasks, Related to Figure 4

Distribution of firing rate scores that compares passes in which the previous or future arms were either a familiar-goal arm or another novel-goal arm, for reference (A) and working (B) memory tasks. Rate scores were calculated by the difference of the rates divided by their sums, so that a positive rate score reflects higher rates during novel-novel passes, while negative rate scores denote a stronger firing in passes that combine familiar- and novel-goal arms. The inset illustrates the meaning of the rate scores on the histograms. The thicker arrows indicate which arm combinations of a trajectory exhibit higher rates on the maze segment highlighted in grey for negative or positive scores. N=novel arm F=Familiar arm. Rate score distributions were calculated in the first half (grey) and second half (back) of the learning trial. Top curves show all cells while the bottom ones represent ones with significant, speed-compensated rate modulation (ANCOVA $p<0.05$). Binomial tests: Reference, top histograms: reward-bound first half, n=310, P=0.909, second half, n=310, P<0.001, center-bound first half, n=318, P=0.753, second half, n=318, P<0.001. Bottom histograms: reward-bound first half, n=143, P=0.071, second half, n=143, P<0.01, center-bound: first half, n=127, P=0.255, second half, n=127, P<0.001. Working, top histograms: inbound first half, n=299, P=0.644, second half, n=299, P=0.860, center-bound first half n=334, P=0.568, second half n=334, P=0.741. Bottom histograms: reward-bound first half, n=109, P=0.722, second half, n=109, P=665, center-bound first half, n=126, P=0.584, second half, n=126, P=0.769.

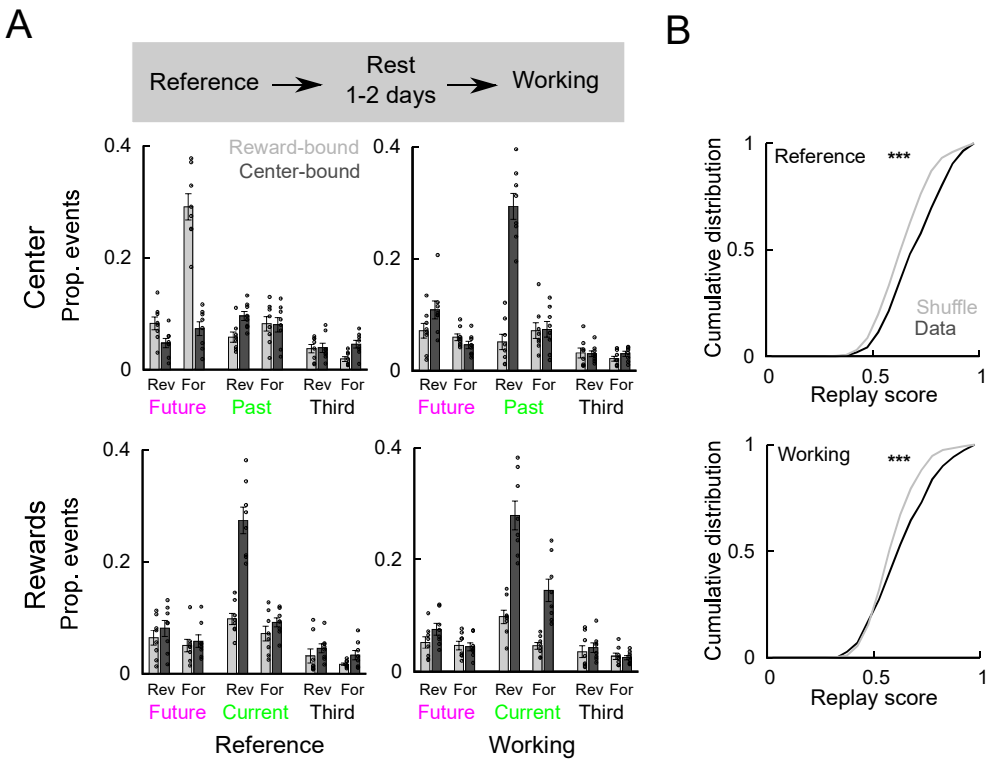


Figure S5. Reactivated trajectories in the working and reference memory tasks when the animal learned and performed the reference memory task first, Related to Figure 7

(A) In three animals, the reference memory task was performed first followed by the working memory task. Proportion (mean ±SEM and individual session values, n=8 sessions) of HSEs in which forward and reverse replay was detected representing future or past/current trajectories of reward-bound vs. center-bound passes. Note that as in the original experiments reference memory HSEs in the center area predicted the future arm choice of the animal in the form of a forward replay of reward-bound trajectories while in the rest of the cases the previous/current arm choices were predicted, as a reverse replay of center-bound trajectories. Three-way ANOVA with Tukey Post hoc, (top left) Reference memory: Future-Forward-Reward-bound group is significantly different from other groups, all $P<0.0000001$; Working memory (top right): Past-Reverse-Center-bound group was significantly different compared to other groups, all $P<0.0000001$; (Bottom). Both tasks: Current-Reverse-Center-bound group vs. other groups, all $P<0.0000001$. (B) The cumulative distribution of replay scores for the original events and those of the shuffled events ($P<0.0000001$, one-tailed K-S test. Reference memory: original n= 446, shuffle n=44600; Working memory: original n=397, shuffle n=39700).

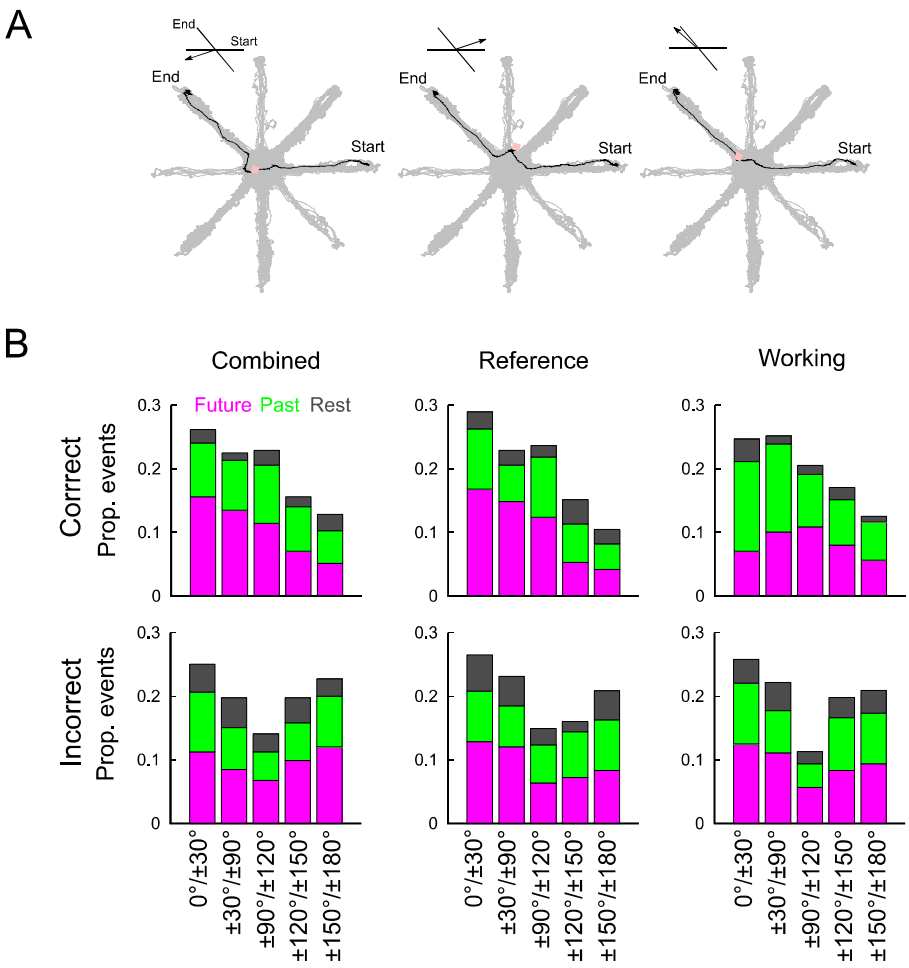


Figure S6. The head direction of the animal at reactivation HSEs, Related to Figure 7

(A) Three examples illustrating the trajectory of the animal passing from one arm to the next. The location of the HSE is marked by the pink circle and arrows (top-left inset) indicate the head direction of the animal. (B) Frequency histograms show how often HSEs occurred in the center area of the maze at different head directions relative to the direction of the goal arm. Both correct trials and error passes are shown. The different color segments show the proportion of events predicting the future and past choice of the animal and the reactivation of the remaining arms.

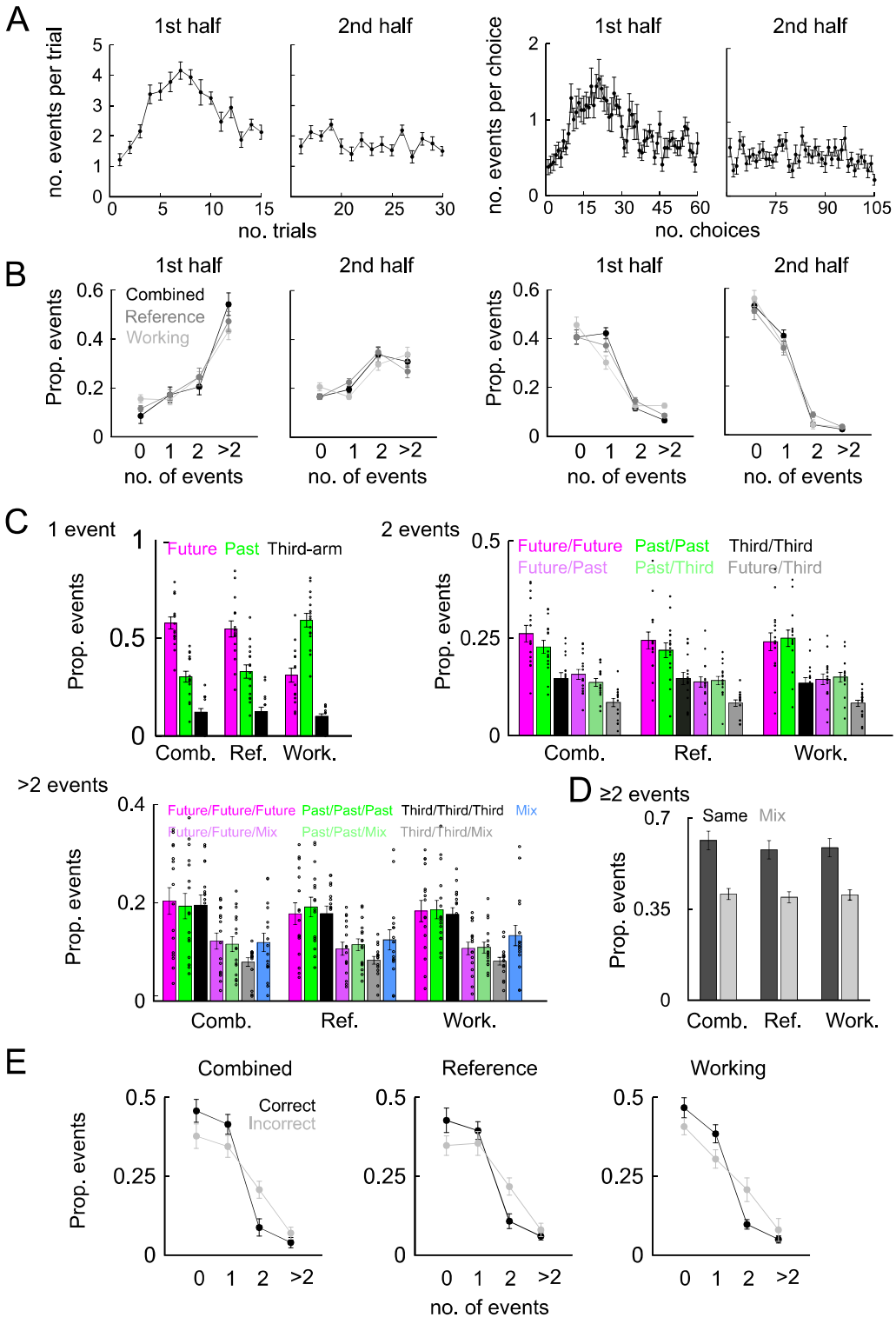

A
1st half
2nd half
no. events per trial
no. trials
no. events per choice
no. choices
B
Combined
Reference
Working
Prop. events
no. of events
C
1 event
Future Past Third-arm
2 events
Future/Future Past/Past Third/Third
Future/Past Past/Third Future/Third
>2 events
Future/Future/Future Past/Past/Past Third/Third/Third Mix
Future/Future/Mix Past/Past/Mix Third/Third/Mix
Comb. Ref. Work.
D ≥2 events
Same Mix
E
Correct
Incorrect

Figure S7. **Number of HSEs and their content within single trials and choice events in the center of the maze,** Related to Figure 7

(A) The average number of HSEs (±SEM, n=16) during different trials (left) and choice passes (right) in the center area of the maze during the course of the learning of the combined task. (B) Distribution of the number of HSEs that occurred within a single trial (left) and within a single choice pass (right). Distribution was calculated separately for the beginning and end of learning (mean ±SEM and individual session values, n=8 sessions). (C) Reactivation content of choice events with single, two or more than two HSEs. Single HSEs during the reference memory tasks predicted the next arm most often, while the past arm dominated during the working memory task (all $P<0.0001$ ANOVA). During choices with multiple replay events in the center (two or more), replay more often reflected the same arm in each event than mixed events (all $P<0.05$ ANOVA with Tukey Post hoc). (D) The proportion of HSEs encoding the same arm or different arms (mixed) within choice events in which two or more HSEs occurred (all $P<0.01554$ ANOVA) (E) Distribution choice events with a different number of HSEs during correct choices and errors. The proportion of multiple (≥2 events) HSEs were significantly higher during error trials (all $P < 0.02$, 2-way ANOVA with Tukey Post hoc).

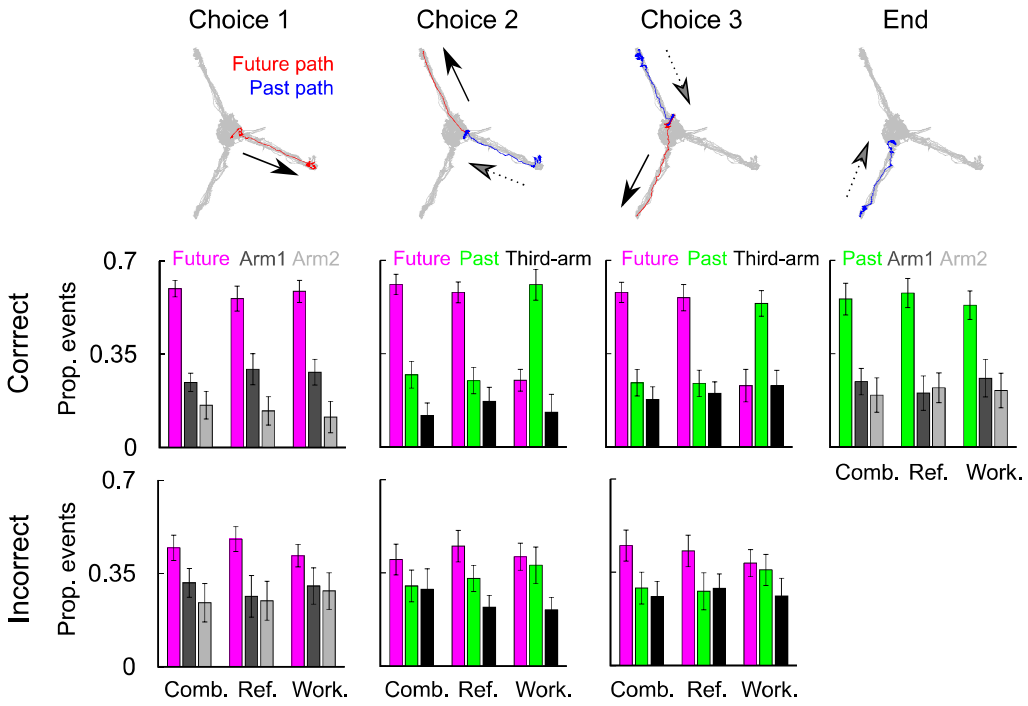


Figure S8. Reactivation content of HSEs at different trial stages associated with different levels of working memory demand, Related to Figure 8

From left to right, choices one to three are shown as well the very end of the trial when the animal returned from the last arm. Top examples display of the animal's trajectory at different trial stages. Middle and bottom histograms show the reactivation content distribution at correct and error passes respectively. During correct passes, in the first choice, replay consistently predicted the future arm, in all three tasks. The future arm is also predicted in the combined and reference memory tasks, during Choice 2 and Choice 3 (all $P<0.0001$ ANOVA Tukey Post hoc), while in the working memory task the past arm was preferentially reactivated (all $P<0.0001$ ANOVA Tukey Post hoc). At the end of the trial, the last arm was preferentially reactivated in all three tasks (all $P<0.0001$ ANOVA Tukey Post hoc). In error passes during Choice 1 in the reference memory task and Choice3 in the combined task the future arm was preferentially reactivated (all $P<0.01$ ANOVA Tukey Post hoc) but in all other cases no arm was preferentially reactivated (all $P>0.095$, ANOVA Tukey Post hoc). In Choice 1 the animal's immediate choice arm was compared with those of the upcoming next two choices whereas in the end the last choice was compared with preceding two choices. Errors cannot be displayed at the end of the trial because a trial ended when the animal collected the last reward, hence, the last choice could not be an error.